On the Rothschild-Stein lifting theorem and its applications

Marco Bramanti
Dipartimento di Matematica, Politecnico di Milano

May 12, 2023

Abstract

This is the text prepared for a minicourse held at Ghent University (Belgium) on January 8-10, 2020, on the topic of Rothschild-Stein lifting and approximation theorem, originally contained in [44]. Besides the original papers where the results discussed here were achieved, which will be quoted in the following, this presentation is based on material taken from [11] and [13].

Contents

1 Context and prerequisites **2**
- 1.1 Hörmander operators . . . 2
- 1.2 A priori estimates . . . 6
- 1.3 Carnot groups . . . 7
- 1.4 Sublaplacians on Carnot groups and their fundamental solutions 11

2 Rothschild-Stein's 1976 paper **16**
- 2.1 The problem, and how to approach it . . . 16
- 2.2 Lifting . . . 18
- 2.3 Approximation with left invariant vector fields . . . 20
- 2.4 Parametrix and L^p estimates . . . 22
- 2.5 Singular integral estimates . . . 26
- 2.6 Final comments on a-priori estimates . . . 27

3 Further applications of the lifting theorem, and variations on the theme **29**
- 3.1 First alternative proofs and applications of the lifting theorem (1970s-1980s) . . . 29
  - 3.1.1 The control distance . . . 30
  - 3.1.2 Geometry of control balls and size estimates on the fundamental solution . . . 31
  - 3.1.3 Jerison's Poincaré inequality for Hörmander's vector fields 38

3.2 More recent variations on the theme and applications of the lifting theorem (2000s-today) 40
3.2.1 A weighted version of the lifting theorem 40
3.2.2 Another version of the lifting theorem for groups 40
3.2.3 Nonsmooth Hörmander's vector fields 42
3.2.4 Homogeneous Hörmander operators 44

4 Sketch of the proof of the lifting and approximation theorem 45
4.1 Lifting 46
4.2 Approximation 53

In these lessons I will address the topic of Rothschild-Stein's "lifting and approximation" theorem. This is an important result contained in the famous paper [44], 1976, by Linda Rothschild and Elias Stein on Acta Mathematica. In that deep paper, lifting and approximation are just one of the techniques used to solve a major problem, related to the proof of "natural" a priori estimates for Hörmander operators. Therefore, an introduction to the lifting theorem should begin with some background about Hörmander operators, and a summary of the main researches that were carried out in the decade spanning from the cornerstone 1967 paper [33] by Lars Hörmander until Rothschild-Stein's paper. This is the content of lesson 1. In lesson 2 I will give a general overview, without proofs, of Rothschild-Stein's paper, dealing not only with the lifting and approximation result, but also with the scope for which this was originally designed. In lesson 3 I will discuss some further developments of the theory of Hörmander operators which are somehow related to lifting and approximation techniques: both direct applications of Rothschild-Stein's results, and extensions of that result, motivated by some further researches. This exposition will not be exhaustive of the subject: I will just choose some topics, reflecting my research interests. Finally, in lesson 4, I will describe the general strategy and some steps of the proof of the lifting theorem, following the alternative approach given by Hörmander-Melin in [34], 1978.

Besides the original papers where the results discussed here were achieved, which will be quoted in the following, this presentation is based on much material taken from [11] and [13].

Acknowledgements. *I wish to thank Michael Ruzhansky for his kind invitation to Ghent University and all the people of Ghent Analysis & PDE group for their keen attention, stimulating conversations, warm hospitality, and delicious cakes.*

1 Context and prerequisites

1.1 Hörmander operators

Hörmander operators are an important class of linear elliptic-parabolic degenerate partial differential operators with smooth coefficients, which have been intensively studied since the late 1960's and are still an active field of research.

To give a brief introduction to this context, let us start recalling the definition of *hypoelliptic operator*.

Definition 1.1 *A differential operator L with $C^{\infty}(\Omega)$ coefficients (Ω open subset of $\mathbb{R}^n$) is said hypoelliptic in Ω if, for any open set $\Omega' \subset \Omega$ and any distribution $u \in \mathcal{D}'(\Omega')$, $Lu \in C^{\infty}(\Omega') \Rightarrow u \in C^{\infty}(\Omega')$.*

A famous result due to Hörmander, [33] 1967, provides an almost complete characterization of *second order hypoelliptic operators with real coefficients*:

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x)\,\partial^2_{x_i x_j} u + \sum_{k=1}^{n} b_k(x)\,\partial_{x_k} u + c(x)\,u$$

Hörmander's preliminary analysis consists in proving that every hypoelliptic operator of this kind has necessarily semi-definite principal part. Now, a second order differential operator with semi-definite principal part can be strongly degenerate, and degenerate operators, in general, are not regularizing. For instance, a solution of $u_{xx} = 0$ in $\mathbb{R}^2$ can be obviously discontinuous. In contrast with this situation, already in 1934, Andrej Kolmogorov [38] exhibited an example of operator of this type, namely

$$Lu = u_{xx} - xu_y - u_t \text{ in } \mathbb{R}^3$$

which, despite its degeneracy, possesses a fundamental solution Γ smooth outside the pole, this fact implying the hypoellipticity of L. The importance of Hörmander's result consists in explaining the origin of these different behaviours. First, Hörmander proves that if L is any linear second order operator with nonnegative characteristic form, then in any open set where the rank of the matrix $\{a_{ij}(x)\}$ is constant, the operator (or its opposite) can be rewritten in the form

$$\mathcal{L} = \sum_{j=1}^{q} X_j^2 + X_0 + c \tag{1.1}$$

where $X_0, X_1, \ldots, X_q$ are real smooth vector fields (that is, first order differential operators):

$$X_j = \sum_{k=1}^{n} b_{jk}(x)\,\partial_{x_k}$$

and c is a smooth function. Given two vector fields X, Y, one can define their *commutator*

$$[X, Y] = XY - YX,$$

which is still a vector field. Let us consider for instance Kolmogorov' operator in $\mathbb{R}^3$

$$Lu = u_{xx} - xu_y - u_t = X_1^2 + X_0$$

with

$$\begin{aligned} X_1 &= \partial_x \\ X_0 &= -x\partial_y - \partial_t. \end{aligned}$$

Note that

$$[X_1, X_0] = -\partial_y.$$

Hence, although the operator L is written in a way that involves only 2 indipendent directions in $\mathbb{R}^3$ (and therefore is degenerate), we see that the missing direction is in some sense recovered by the commutator: the *three* vector fields

$$X_1 = \partial_x; X_0 = -x\partial_y - \partial_t; [X_1, X_0] = -\partial_y.$$

span $\mathbb{R}^3$ at every point.

More generally, we can consider the vector space of all the vector fields on some open set $\Omega \subset \mathbb{R}^n$. Defining the Lie bracket of any two vector fields X, Y as

$$[X, Y] = XY - YX,$$

this structure becomes a *Lie algebra*. Let us recall the standard definition of this abstract structure:

Definition 1.2 *A* Lie algebra *(over $\mathbb{R}$) is a real vector space $(\mathfrak{g}, +, \cdot)$ endowed with another internal operation, called* Lie bracket*,*

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

enjoying the following properties:

bilinearity:

$$[\lambda X + \mu Y, Z] = \lambda [X, Z] + \mu [Y, Z]$$

anticommutativity:

$$[X, Y] = -[Y, X];$$

Jacobi identity:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

for every $X, Y, Z \in \mathfrak{g}$, $\lambda, \mu \in \mathbb{R}$.

Given a family of vector fields $X_0, X_1, ..., X_q$ we can consider the Lie (sub)algebra generated by them, $\mathrm{Lie}(X_0, X_1, ..., X_q)$, that is, the smallest Lie algebra containing $X_0, X_1, ..., X_q$. Roughly speaking, $\mathrm{Lie}(X_0, X_1, ..., X_q)$ consists in linear combination of the vector fields X_0, X_1,..., X_q and their (iterated) commutators.

Evaluating the vector fields of this algebra at some point of Ω we find a vector subspace of $\mathbb{R}^n$. Then:

Definition 1.3 *We say that a system of (real, smooth) vector fields $X_0, X_1, ..., X_q$ defined in some open set $\Omega \subset \mathbb{R}^n$ satisfies Hörmander's condition in Ω if the Lie algebra generated by $X_0, X_1, \ldots, X_q$ span the whole $\mathbb{R}^n$ at every point $x \in \Omega$:*

$$\dim \{X_x : X \in \mathrm{Lie}(X_0, X_1, ..., X_q)\} = n \quad \forall x \in \Omega.$$

We also say that $X_0, X_1, ..., X_q$ are a system of Hörmander's vector fields.

In this case, we say that Hörmander's condition holds at step s $(= 2, 3, ...)$ if s is the maximum length of commutators required to satisfy the above condition in Ω.

Hörmander's theorem then reads as follows:

Theorem 1.4 (Hörmander, [33]) *Let $X_0, X_1, ..., X_q$ be real smooth vector fields in some open set $\Omega \subseteq \mathbb{R}^n$, satisfying Hörmander's condition in Ω and let $c \in C^\infty(\Omega)$. Then the operator*

$$L = \sum_{j=1}^{q} X_j^2 + X_0 + c,$$

is hypoelliptic in Ω.

We will say that such L is a *Hörmander operator.*

Conversely, if in an open set $U \subset \Omega$ the rank of the Lie algebra (that is, the dimension of the vector space spanned by iterated commutators) is constant and strictly less than n, then the operator $\mathcal{L}$ is not hypoelliptic in U. Hence Hörmander's condition is "almost necessary" for hypoellipticity.

For a discussion of some motivations to study operators of type (1.1), the reader is referred for instance to [11, Chap. 2].

The vector field X_0 is sometimes called *drift*. It is important to distinguish between two situations:

(1). X_0 is required to satisfy Hörmander's condition in Ω, that is the set $X_1, \ldots, X_q$ alone does not fulfill the condition. In this case the drift is essential for hypoellipticity, like the time derivative is essential for the validity of the good properties of the heat operators.

(2). The set $X_1, \ldots, X_q$ fulfills Hörmander's condition. Then the vector field X_0, if present, plays the role of a lower order term, and is not essential. In this situation the prototype operator to be studied is

$$\mathcal{L} = \sum_{j=1}^{q} X_j^2,$$

which is known as "sum of squares of Hörmander's vector fields" or "sublaplacian".

Example 1.5 *The simplest example of the situation (2) is the* sublaplacian on the Heisenberg group $\mathbb{H}^1$*:*

$$L = X^2 + Y^2 \equiv (\partial_x + 2y\partial_t)^2 + (\partial_y - 2x\partial_t)^2 \tag{1.2}$$

defined in $\mathbb{R}^3$. Note that

$$\begin{aligned} X &= \partial_x + 2y\partial_t \\ Y &= \partial_y - 2x\partial_t \\ [X, Y] &= -4\partial_t \end{aligned}$$

span $\mathbb{R}^3$ at every point: we can say that X, Y satisfy Hörmander's condition at step 2, and by Hörmander's theorem the operator L is hypoelliptic in $\mathbb{R}^3$.

Throughout these lessons we will concentrate on Hörmander operators of the kind "sum of squares", for the seak of simplicity. Actually, the study of Hörmander operators with drift poses substantially harder problems.

1.2 A priori estimates

A key point in the proof of Hörmander's theorem [33], which has an independent interest, and has been particularly stressed in the alternative proof given some years later by Kohn [37] and independently by Oleĭnik-Radkevič [43], consists in establishing the so-called *subelliptic estimates*. These estimates can be proved by techniques of pseudodifferntial operators, and involve the standard Sobolev spaces $H^s(\mathbb{R}^n)$ of fractional order, defined via Fourier transform: for every $s \in \mathbb{R}$ we let

$$\|u\|_{H^s(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \left(1 + |\xi|^2\right)^s d\xi \right)^{1/2}$$

The precise result reads as follows:

Theorem 1.6 (Subelliptic estimates) *Let L be a Hörmander operator. There exists $\varepsilon \in (0,1)$ such that for every couple of cutoff functions $\eta, \eta' \in C_0^\infty(\Omega)$ satisfying $\eta \prec \eta'$ (that is, $\eta' = 1$ on $\operatorname{supp} \eta$) and for every $\sigma, \tau \geq 0$, there exists a constant $C > 0$, such that*

$$\|\eta u\|_{H^{\sigma+\varepsilon}} \leqslant C \left(\|\eta' L u\|_{H^\sigma} + \|\eta' u\|_{H^{-\tau}} \right)$$

whenever $u \in \mathcal{D}'(\Omega)$ is such that the right hand side of the above inequality is finite.

Note that for every distribution u and cutoff function η', there exists some $\tau > 0$ such that $\|\eta' u\|_{H^{-\tau}}$ is finite. Therefore the previous estimate implies that if Lu is smooth also u is smooth. In the special case of a "sum of squares" operator, one can take $\varepsilon = 1/s$ where s is step of Hörmander's condition. In the elliptic case, we would have $s = 1 = \varepsilon$, whence the term "subelliptic estimates", since here in general $\varepsilon < 1$.

For a survey on and a soft introduction to the proof of Hörmander theorem and subelliptic estimates, see [12].

Subelliptic estimates open a natural problem. We are able to bound just a fractional derivative of u in terms of Lu, even though the operator L is highly regularizing. This is somehow unsatisfactory. To put it into another way: we know that $Lu \in C^\infty(\Omega)$ implies $u \in C^\infty(\Omega)$ but if Lu has some partial regularity, for instance $Lu \in H^{k,2}$ for *some* (but not for all) k, or if $Lu \in L^p(\Omega)$ for some $p \neq 2$, subelliptic estimates do not allow us to deduce the natural gain of regularity of u. The point is that we are trying to bound the usual, "Cartesian" derivatives for an operator which is highly anisotropic: we can expect L

to control the specific directions given by the vector fields X_i. So perhaps the situation could be better if we tried to bound the derivatives $X_j u$ and $X_i X_j u$. But this requires a completely different approach: the techniques of pseudodifferential operators used to get subelliptic estimates are shaped on *fractional* but *isotropic* derivatives; moreover, they privilege L^2 bounds with respect to L^p bounds for other p's; also, if one tries to get a bound on the derivatives with respect to the vector fields, the Hilbert space technique will offer an estimate on *first order derivatives*, like in the following "energy estimate" that we can prove for the sublaplacian on $\mathbb{H}^1$ (see Example 1.5):

$$\|Xu\|^2 + \|Yu\|^2 \leq \left| \int Lu \cdot u \right| \leq \|Lu\| \cdot \|u\| \lesssim \|Lu\|^2 + \|u\|^2 .$$

Instead, if one wants to prove an L^p bound on $X_i X_j u$ (and $X_0 u$, if the drift is present) in terms of Lu and u, then one should mimic the techniques used to prove L^p estimates for *strong solutions to nonvariational elliptic equations.* This involves the use of representation formulas by means of fundamental solutions, and the application of singular integral estimates (Calderón-Zygmund theory). This program has been actually carried out, for Hörmander operators, in three famous, outstanding papers of the mid-seventies:

1974, Folland-Stein, Comm. Pure Appl. Math., [31]

1975, Folland, Arkiv für Mat., [27]

1976, Rothschild-Stein, Acta Math., [44]

In these papers a priori estimates of the kind

$$\|X_i X_j u\|_p + \|X_0 u\|_p \lesssim \|Lu\|_p + \|u\|_p \quad \text{for } 1 < p < \infty$$

have been proved, at increasing levels of generality. Namely:

Step 1: 1974, Folland-Stein: the Kohn-Laplacian on Heisenberg groups (part I) and on nondegenerate CR manifolds (part II);

Step 2: 1975, Folland: sublaplacians on homogeneous groups;

Step 3: 1976, Rothschild-Stein: general Hörmander operators.

These papers introduced a number of fundamental ideas which, still now, represent some of the basic tools which are necessary to do research in this field.

We want to stress the fact that each further step does not make the previous ones "useless". Namely: the results of Step 3 extend those of Step 2, but are also based on them; the main result of part I of Step 1 is in some sense properly contained in Step 2; however, part II of Step 1 is not contained in Step 2; it consists in an extension of the results of Part I to a more general situation, and that way of reasoning can be considered as the seed of the ideas used in Step 3 to extend Step 2.

I suggest to the interested reader the survey paper [29] by Folland, 1977, which contains a good introduction to these three papers.

1.3 Carnot groups

Let us start recalling some basic definitions and facts about Carnot groups. For this material and for the proofs of some facts we refer to the original paper [27] or to Stein's book [46, Chap.XIII, §5]; a much wider presentation of this theory can be found in the monograph by Bonfiglioli, Lanconelli, Uguzzoni [5]; this last book is particularly suggested to find a good wealth of explicit examples of homogeneous groups and corresponding sublaplacians, with computations carried out in detail.

A *homogeneous group* (in $\mathbb{R}^N$) is a Lie group $\left(\mathbb{R}^N, \circ\right)$ (where the group operation $\circ$ will be thought as a "translation", and 0 is the identity group) endowed with a one parameter family $\{D_\lambda\}_{\lambda>0}$ of group automorphisms ("dilations") which act this way:

$$D_\lambda\left(x_1, x_2, ..., x_N\right) = \left(\lambda^{\alpha_1} x_1, \lambda^{\alpha_2} x_2, ..., \lambda^{\alpha_N} x_N\right) \tag{1.3}$$

for suitable integers $1 = \alpha_1 \leqslant \alpha_2 \leqslant ... \leqslant \alpha_N$. We will write $\mathbb{G} = \left(\mathbb{R}^N, \circ, D_\lambda\right)$ to denote this structure. The number

$$Q = \sum_{i=1}^{N} \alpha_i$$

will be called *homogeneous dimension* of $\mathbb{G}$. Under the change of coordinates $x = D_\lambda\left(y\right)$ the volume element transforms according to

$$dx = \lambda^Q dy \tag{1.4}$$

which justifies the name of homogeneous dimension for Q. Note that we always have $Q \geqslant N$, and $Q = N$ only if the dilations are the Euclidean ones.

In a homogeneous group the volume element is also invariant with respect to left translation, right translation, and inversion. Moreover, it is always possible to choose a system of coordinates in $\mathbb{G}$ such that $u^{-1} = -u$.

Example 1.7 *The Heisenberg group $\mathbb{H}^n$ in $\mathbb{R}^{2n+1}$. Writing the points of $\mathbb{R}^{2n+1}$ as*

$$(x, y, t) \equiv (x_1, x_2, ..., x_n, y_1, y_2, ..., y_n, t) \in \mathbb{R}^{2n+1}$$

we can express the group law of $\mathbb{H}^n$ as:

$$(x, y, t) \circ (x', y', t') = \left(x + x', y + y', t + t' + 2\left(y \cdot x' - x \cdot y'\right)\right)$$

and the dilations as

$$D\left(\lambda\right)(x, y, t) = \left(\lambda x, \lambda y, \lambda^2 t\right).$$

The set $\mathbb{R}^{2n+1}$ with this group law is called the Heisenberg group $\mathbb{H}^n$.

A *homogeneous norm* on $\mathbb{G}$ (also called a *gauge* on $\mathbb{G}$) is a continuous function

$$\|\cdot\| : \mathbb{R}^N \to [0, +\infty),$$

such that, for some constant $c > 0$, for every $x, y \in \mathbb{R}^N$,

$$\begin{array}{ll} (i) & \|x\| = 0 \Longleftrightarrow x = 0 \\ (ii) & \left\|D_\lambda\left(x\right)\right\| = \lambda\left\|x\right\| \;\; \forall \lambda > 0 \\ (iii) & \left\|x^{-1}\right\| \leqslant c\left\|x\right\| \\ (iv) & \left\|x \circ y\right\| \leqslant c\left(\|x\| + \|y\|\right). \end{array}$$

We will always use the symbol $\|\cdot\|$ to denote a homogeneous norm, and the symbol $|\cdot|$ to denote the Euclidean norm.

Concrete ways to define a homogeneous norm on $\mathbb{G}$ are for instance the following:

$$\|x\| = \max_{k=1,2,\dots,N} |x_k|^{\frac{1}{\alpha_k}}$$

or

$$\|x\| = \left(\sum_{k=1}^{N} |x_k|^{\frac{Q}{\alpha_k}} \right)^{1/Q}.$$

The following property can be proved quite easily, and can be helpful to check that some explicit function is actually a homogeneous norm:

Proposition 1.8 *Let*

$$\|\cdot\| : \mathbb{R}^N \to [0, +\infty),$$

be a continuous function satisfying conditions (i)-(ii) in the above definition. Then it also satisfies (iii)-(iv). Moreover, the sets

$$B_R = \left\{ x \in \mathbb{R}^N : \|x\| \le R \right\}$$

are compact in the Euclidean sense.

We say that a smooth function f in $\mathbb{R}^N \setminus \{0\}$ is *D_λ-homogeneous of degree* $\beta \in \mathbb{R}$ (or simply "β-homogeneous") if

$$f(D_\lambda(x)) = \lambda^\beta f(x) \quad \forall \lambda > 0, x \in \mathbb{R}^N \setminus \{0\}.$$

Given any differential operator P with smooth coefficients on $\mathbb{R}^N$, we say that P is *left invariant* if for every $x, y \in \mathbb{R}^N$

$$P(L_y f)(x) = L_y(Pf(x))$$

for every smooth function f, where

$$L_y f(x) = f(y \circ x).$$

Analogously one defines the notion of *right invariant* differential operator. Also, P is *β-homogeneous* (for some $\beta \in \mathbb{R}$) if

$$P(f(D_\lambda(x))) = \lambda^\beta (Pf)(D_\lambda(x))$$

for every smooth function f, $\lambda > 0$ and $x \in \mathbb{R}^N$.

Clearly, if P is a differential operator homogeneous of degree δ_1 and f is a homogeneous function of degree δ_2, then Pf is homogeneous of degree $\delta_2 - \delta_1$. For example, $x_i \frac{\partial}{\partial x_j}$ is homogeneous of degree $\alpha_j - \alpha_i$.

Within the Lie algebra of all the smooth vector fields on $\mathbb{R}^N$ we can consider the *Lie algebra* $\mathfrak{g}$ *of left invariant vector fields* over $\mathbb{G}$. This Lie algebra is readily seen to be finite dimensional. Namely, the *canonical base* of $\mathfrak{g}$ consists in the

N vector fields $X_1, X_2, \dots, X_N$, where each X_i is, by definition, the only left invariant vector field that agrees with ∂_{x_i} at the origin.

We assume that for some integer $q < N$ the vector fields $X_1, X_2, \dots, X_q$ are 1-homogeneous and the Lie algebra generated by them is $\mathfrak{g}$. If s is the maximum length of commutators

$$\left[X_{i_1}, \left[X_{i_2}, ..., \left[X_{i_{s-1}}, X_{i_s}\right]\right]\right], \quad i_j \in \{1, 2, ..., q\}$$

required to span $\mathfrak{g}$, then we will say that $\mathfrak{g}$ is a *stratified Lie algebra of step* s, $\mathbb{G}$ is a *Carnot group* (or a stratified homogeneous group) and the *generators* $X_1, X_2, \dots, X_q$ satisfy Hörmander's condition at step s in $\mathbb{R}^N$.

Under these assumptions, let us denote by

$$\mathcal{L} = \sum_{j=1}^{q} X_j^2$$

the *canonical sublaplacian* on $\mathbb{G}$. This operator is left invariant, 2-homogeneous (with respect to the dilations on $\mathbb{G}$) and, by Hörmander's theorem, hypoelliptic in $\mathbb{R}^N$.

For homogeneity reasons, the vector fields X_j have a triangular structure

$$X_j = \partial_{x_j} + \sum_{k=j+1}^{N} b_k\left(x_1, ..., x_j\right) \partial_{x_k}, \quad j = 1, 2, ..., N$$

with b_k homogeneous polynomials, which in particular implies that, for every $\phi \in C_0^1\left(\mathbb{R}^N\right)$, $u \in C^1\left(\mathbb{R}^N\right)$, one has

$$\int_{\mathbb{R}^N} \left(X_j u\right) \phi = -\int_{\mathbb{R}^N} u \left(X_j \phi\right).$$

In other words, for the *adjoint* X_j^* we simply have

$$X_j^* = -X_j, j = 1, 2, ..., N$$

and in particular

$$\mathcal{L}^* = \mathcal{L}.$$

Example 1.9 *On the Heisenberg group* $\mathbb{H}^n$ *(see Example 1.7) the standard notation for the generators is:*

$$\begin{aligned} X_j &= \partial_{x_j} + 2y_j \partial_t \\ Y_j &= \partial_{y_j} - 2x_j \partial_t. \end{aligned}$$

Note that

$$\left[X_j, Y_j\right] = -4\partial_t$$

for every $j = 1, 2, ..., n$*. The sublaplacian is*

$$L = \sum_{j=1}^{n} \left(X_j^2 + Y_j^2\right)$$

and generalizes the operator in Example 1.5.

We will make use of the *Sobolev spaces* $W_X^{k,p}(\mathbb{G})$ *induced by the system of vector fields*

$$X = \{X_1, X_2, \ldots, X_q\}.$$

More precisely, given an open subset Ω of $\mathbb{R}^N$, we say that $f \in W_X^{1,p}(\Omega)$ $(1 \leqslant p \leqslant \infty)$ if $f \in L^p(\Omega)$ and there exist, in weak sense, $X_j f \in L^p(\Omega)$ for $j = 1, 2, ..., q$. Inductively, we say that $f \in W_X^{k,p}(\Omega)$ for $k = 2, 3, ...$ if $f \in W_X^{k-1,p}(\Omega)$ and any weak derivative of order $k-1$ of f, $X_{j_1} X_{j_2} ... X_{j_{k-1}} f$, belongs to $W_X^{1,p}(\Omega)$. We set

$$\|f\|_{W_X^{k,p}(\Omega)} = \|f\|_{L^p(\Omega)} + \sum_{h=1}^{k} \sum_{j_i = 1,2,\ldots,q} \|X_{j_1} X_{j_2} ... X_{j_h} f\|_{L^p(\Omega)}.$$

The *convolution* of two functions in $\mathbb{G}$ is defined as

$$(f * g)(x) = \int_{\mathbb{R}^N} f(x \circ y^{-1})\, g(y)\, dy = \int_{\mathbb{R}^N} g(y^{-1} \circ x)\, f(y)\, dy, \tag{1.5}$$

for every couple of functions for which the above integrals make sense. Note that this convolution *is not commutative.* If P is any *left invariant differential operator*,

$$P(f * g) = f * Pg \tag{1.6}$$

(provided the integrals converge). Note that, differently from the Euclidean case, we cannot interchangeably take the differential operator P on f or g. Observe that, like P is left invariant with respect to translation, (1.6) says that it is left invariant with respect to convolution. (Actually, to remember the positions of the variables in definition (1.5), one can think that it is given this way just in order for (1.6) to be true).

1.4 Sublaplacians on Carnot groups and their fundamental solutions

Let us now sketch some of the main ideas contained in Folland's famous paper [27], 1975. A fundamental result proved in that paper is the following:

Theorem 1.10 (Existence of a homogeneous fundamental solution) *Let $\mathcal{L}$ be a left invariant differential operator homogeneous of degree two on a homogeneous group $\mathbb{G}$, such that $\mathcal{L}$ and $\mathcal{L}^*$ are both hypoelliptic. Moreover, assume* $Q \geq 3$ *(where Q is the homogeneous dimension of $\mathbb{G}$). Then there is a unique translation invariant fundamental solution Γ (that is, $\Gamma(x,y) = \Gamma\left(y^{-1} \circ x\right)$) such that:*

(*a*) $\Gamma \in C^\infty\left(\mathbb{R}^N \setminus \{0\}\right)$;
(*b*) Γ *is homogeneous of degree* $(2 - Q)$;
(*c*) *for every* $u \in C_0^\infty\left(\mathbb{R}^N\right)$

$$u(x) = \int_{\mathbb{R}^N} \Gamma(y^{-1} \circ x) \mathcal{L} u(y)\, dy = \mathcal{L} \int_{\mathbb{R}^N} \Gamma(y^{-1} \circ x) u(y)\, dy \tag{1.7}$$

(actually, the identities hold for every distribution u*, provided we properly interpret convolutions).*

The result proved by Folland is actually more general: he assumes $\mathcal{L}$ homogeneous of degree $\alpha > 0$ and $Q > \alpha$. We will not be interested in the case $\alpha \neq 2$. Note that the restriction $Q > 2$ is irrelevant, since the only case it excludes is that of elliptic equations in two variables with constant coefficients. This theorem in particular applies to the sublaplacian on $\mathbb{G}$ (for which $\mathcal{L} = \mathcal{L}^*$ is hypoelliptic, left invariant and 2-homogeneous) and from now on, for the sake of simplicity we will restrict to the case of sublaplacians. However, note that in the statement of the theorem nothing is required about the *explicit form* of the operator $\mathcal{L}$.

Let us explicitly note the relevance of the information contained in the *two* identities (1.7). Namely, saying that Γ is a fundamental solution for $\mathcal{L}$ means that

$$\begin{gathered}\mathcal{L}\Gamma = \delta, \text{ which implies} \\ \mathcal{L}(\tau * \Gamma) = \tau * \mathcal{L}\Gamma = \tau * \delta = \tau.\end{gathered}$$

However, due to the noncommutativity of the convolution on $\mathbb{G}$, we cannot write, in general $\mathcal{L}\,(f * g) = \mathcal{L}f * g$, hence the identity $\mathcal{L}(\tau * \Gamma) = (\mathcal{L}\tau) * \Gamma$ is not trivial, but is actually a further information which is supplied by the theorem, which implies that we can represent any distribution τ as $\tau = (\mathcal{L}\tau) * \Gamma$.

The proof of the above theorem is nonconstructive. First, applying some deep abstract results from distribution theory, Folland shows the existence of a "local fundamental solution", smooth outside the pole by Hörmander's hypoellipticity theorem. Then, starting from this kernel and exploiting the dilations on $\mathbb{G}$ and the homogeneity of $\mathcal{L}$, he builds a new, globally defined, kernel which also possesses the desired homogeneity. In the end, we do not have any idea of the explicit form of this Γ. However, its abstract properties are really enough to make the theory work very well.

First of all, let us note that since Γ is $(2-Q)$-homogeneous and $X_1, ..., X_q$ are 1-homogeneous we have that

$$\begin{gathered}X_i\Gamma \text{ is } (1-Q)\text{-homogeneous} \\ X_iX_j\Gamma \text{ is } (-Q)\text{-homogeneous}\end{gathered}$$

(for $i, j = 1, 2, ..., q$). In particular, Γ and $X_i\Gamma$ are locally integrable, while $X_iX_j\Gamma$ is not.

Secondly, since $X_1, ..., X_q$ are left invariant, for every $u \in C_0^\infty\left(\mathbb{R}^N\right)$, starting with the representation formula (1.7) we can compute

$$X_j u\,(x) = \int_{\mathbb{R}^N} (X_j\Gamma)\,(y^{-1} \circ x)\mathcal{L}u\,(y)\,dy \text{ for } j = 1, 2, ..., q.$$

Now, if we try to compute another derivative, that is

$$X_iX_j u\,(x) = X_i \int_{\mathbb{R}^N} (X_j\Gamma)\,(y^{-1} \circ x)\mathcal{L}u\,(y)\,dy$$

for $i,j=1,2,...,q$, we can no longer take the derivative X_i inside the integral, since $X_iX_j\Gamma$ is not integrable. This is the point where the *theory of singular integrals* cames in. Thanks to the homogeneity properties of Γ and the rich underlying structure of Carnot group, the following result can be proved:

Theorem 1.11 *The singular kernel $K(x)=X_iX_j\Gamma(x)$ satisfies the following properties:*

(i)

$$|K(x)|\leq\frac{c}{\|x\|^Q}\text{ for every }x\neq0;$$

(ii)

$$|K(x\circ y)-K(x)|+|K(y\circ x)-K(x)|\leq c\frac{\|y\|}{\|x\|^{Q+1}}$$

whenever $\|x\|\geq2\|y\|$;

(iii)

$$\int_{R_1<\|x\|<R_2}K(x)\,dx=\int_{R_1<\|x\|<R_2}K\left(x^{-1}\right)dx=0$$

for any $0<R_1<R_2<\infty$.

Moreover, for every $i,j=1,2,...,q$ there exists a constant c_{ij} such that the following representation formula holds, for any test function u:

$$X_iX_ju(x)=\lim_{\varepsilon\to0^+}\int_{\|y^{-1}\circ x\|>\varepsilon}X_iX_j\Gamma(y^{-1}\circ x)(\mathcal{L}u)(y)\,dy+c_{ij}\mathcal{L}u(x).\quad(1.8)$$

Inequalities (i)-(ii) are usually called the *standard estimates of singular integral kernels*, while (iii) is a strong and very useful form of *vanishing property* for K. Thanks to (iii) the *principal value integral* appearing in (1.8) can be proved to exist, even though the integral of $X_iX_j\Gamma$ is not absolutely convergent.

The statements contained in the above theorem are surprisingly similar to the ones that can be proved, on the Euclidean group $\left(\mathbb{R}^N,+\right)$ for the second derivatives of the fundamental solution of the classical laplacian

$$\gamma(x)=\frac{c}{|x|^N}$$

(or any other constant coefficient elliptic operator). In that context, the theory of singular integrals developed by Calderón-Zygmund already in the mid 1950's (see e.g. [21], [23], [22]), provided L^p estimates of the kind

$$\left\|u_{x_ix_j}\right\|_{L^p(\mathbb{R}^N)}\leq c\|\Delta u\|_{L^p(\mathbb{R}^N)}\quad\text{for }i,j=1,2,...,N,1<p<\infty.$$

Now, despite the stiking analogy between the properties of $\gamma_{x_ix_j}$ in the Euclidean case, and the properties of $X_iX_j\Gamma$ in the case of Carnot groups, getting analogous L^p estimates on the second order derivatives $X_i X_j u$ is by no means a straightforward result. Instead, this result required the development of a substantial piece of *abstract theory of singular integrals*, extending the original ideas of Calderón-Zygmund to more general contexts. This is actually another story, that will not be told here. Let us just mention a couple of cornerstones in this area, both dating 1971: the long paper [36] by Knapp-Stein, where the tools necessary to get L^2 estimates for singular integrals in the special context of homogeneous groups were introduced, and the monograph [25] by Coifman-Weiss, where the basic theory of singular integrals in *spaces of homogeneous type* was introduced, allowing in particular to derive L^p estimates from every $p \in (1, \infty)$ from L^2 estimates, once we know that the singular kernel satisfies the standard estimates. By the way, let us recall the by now standard definition of space of homogeneous type:

Definition 1.12 *Let X be a set endowed with a quasidistance, that is a function*

$$d : X \times X \to [0, +\infty)$$

such that

$$\begin{aligned} &d(x, y) = 0 \Leftrightarrow x = y \\ &d(x, y) = d(y, x) \\ &d(x, y) \leq c\,[d(x, z) + d(z, y)] \end{aligned}$$

for some constant $c \geq 1$ and every $x, y, z \in X$. The d-balls $B(x, r)$ induce a topology in X. Assume that the d-balls are open with respect to this topology. Moreover, assume there exists on X a Borel measure μ such that the doubling condition holds:

$$0 < \mu(B(x, 2r)) \leq c\mu(B(x, r)) < \infty$$

for some constant $c > 0$, every $x \in X$ and $r > 0$. Then (X, d, μ) is called a space of homogeneous type.

Combininig the results in Thm 1.11 with the abstract theory of singular integrals developed within the early 1970's, Folland could prove the a priori estimate

$$\|X_i X_j u\|_{L^p(\mathbb{R}^N)} \leq c\,\|Lu\|_{L^p(\mathbb{R}^N)} \quad \text{for } i, j = 1, 2, ..., q,$$

any test function u, $1 < p < \infty$, where $L = \sum_{i=1}^{q} X_i^2$ is a sublaplacian on a Carnot group. With an extra effort, highly nontrivial in this noncommuting context, one can also prove higher order regularity result of the following kind:

Theorem 1.13 (Local Sobolev regularity and solvability) *Let L be the sublaplacian on a Carnot group $\mathbb{G}$ of homogeneous dimension $Q > 2$, let Ω be a bounded domain in $\mathbb{G}$, let $\Omega' \Subset \Omega'' \Subset \Omega$ and let $k \in \mathbb{N}$. If u is a distribution in Ω such that $Lu \in W_X^{k,p}(\Omega)$ $(1 < p < \infty)$ then $u \in W_X^{k+2,p}(\Omega'')$ and*

$$\|u\|_{W_X^{k+2,p}(\Omega')} \leq c\left\{\|Lu\|_{W_X^{k,p}(\Omega)} + \|u\|_{L^p(\Omega'')}\right\}.$$

Moreover, for every $f \in L^p(\Omega)$ there exists $u \in W^{2,p}_{X,loc}(\Omega)$ satisfying the equation $Lu = f$ in Ω (and therefore satisfying the above a priori estimate for $k = 0$).

The following global result also holds:

Theorem 1.14 (Global Sobolev regularity) *Under the above assumptions, for every $k \in \mathbb{N}$ and $1 < p < \infty$ there exists a constant c such that if $u \in L^p\left(\mathbb{R}^N\right)$ and $Lu \in W^{k,p}_X\left(\mathbb{R}^N\right)$, then $u \in W^{k+2,p}_X\left(\mathbb{R}^N\right)$ and*

$$\|u\|_{W^{k+2,p}_X(\mathbb{R}^N)} \leqslant c\left(\|Lu\|_{W^{k,p}_X(\mathbb{R}^N)} + \|u\|_{L^p(\mathbb{R}^N)}\right). \tag{1.9}$$

The above results are proved in [27], although not exactly in this form and with this language. A detailed proof of these results will be written in [13].

Folland's paper [27], whose results have been just summarized, represents a major advance with respect to the previous results contained in the first part of the paper [31] by Folland-Stein, 1974, where similar results are obtained for the sublaplacians on the Heisenberg groups $\mathbb{H}^n$, exploiting the knowledge of an *explicit* fundamental solution (which was known in that case). But, besides the more general class of groups which is considered by Folland, the research program which is clearly stated in [27] consists in thinking that the homogeneous group case could be the starting point to get similar results for general Hörmander operators, *by approximation*:

> "In the theory of *elliptic* operators the constant-coefficient operators serve as a useful class of models for the general situation: constant-coefficient operators are amenable to treatment by the techniques of Euclidean harmonic analysis (Fourier transforms, convolution operators, etc.), and the results obtained thereby can usually be extended to the variable-coefficient case by perturbation arguments. Now, a constant-coefficient operator is nothing more than a translation-invariant operator on the Abelian Lie group $\mathbb{R}^N$. From this point of view, it is natural to attempt to construct a class of models for non-elliptic operators of the sort discussed above among the translation-invariant operators on certain non-Abelian Lie groups. The Lie algebras of the groups involved should have a structure which reflects the behavior of the commutators in the original problem and the groups themselves should admit a "harmonic analysis" which will produce results similar to those of the Euclidean case. A particular case of this program, has been carried out in considerable detail in Folland-Stein [30], [31], in which sharp L^p and Lipschitz (or Hölder) estimates for the $\overline{\partial}_b$ complex on the boundary of a complex domain with nondegenerate Levi form are obtained by using certain left-invariant operators on the Heisenberg group as models".
>
> Folland, [27]

This research program has been fully carried out by Rothschild-Stein, [44], 1976, and in doing so the "lifting and approximation" technique was devised.

2 Rothschild-Stein's 1976 paper

Let us discuss now some ideas contained in the famous paper by Rothschild-Stein, 1976, Acta Math. [44], where the *lifting theorem* firstly appeared.

2.1 The problem, and how to approach it

Here the issue is to prove L^p estimates on $X_i X_j u$ of the same kind already proved on homogeneous groups, for a general Hörmander operator, that is when an underlying group structure is lacking. Let us recall two examples of Hörmander operators which do not admit an underlying homogeneous group:

(i) The Grushin operator:

$$L = \partial_{xx}^2 + x^2 \partial_{yy}^2. \tag{2.1}$$

Here the absence of an underlying group of translation is made evident by the fact that at points (x_0, y_0) with $x_0 \neq 0$ the vector fields $X_1 = \partial_x, X_2 = x\partial_y$ are independent, while if $x_0 = 0$ they are not. In contrast with this, if two left invariant vector fields are independent at one point, they must be independent everywhere. The reason is that if $c_1 X_1 + c_2 X_2 = 0$ at some point, with c_1, c_2 constants, then the vector field $Y = c_1 X_1 + c_2 X_2$ is also left invariant and vanishes at some point; this implies that it vanishes everywhere. (A left invariant vector field is *uniquely determined* by its value at one point).

(ii) The Mumford operator (appearing in computer vision, in the study of the "process of random direction", see [41]):

$$L = \partial_t + \cos\theta \partial_x + \sin\theta \partial_y + \frac{\sigma^2}{2} \partial_{\theta\theta}^2.$$

Here the absence of a family of dilations is made evident by the fact that the coefficients $\cos\theta, \sin\theta$ are not polynomials.

Now, what does it mean to approximate a general Hörmander operator by means of a 2-homogeneous left invariant one, according to Folland's research program quoted above? Let us give a simple example:

Example 2.1 *The vector fields*

$$X = \partial_x + 2y\partial_t; Y = \partial_y - 2(e^x - 1)\partial_t$$

satisfy Hörmander's condition at step 2*; the coefficients of* Y *are not polynomials (hence cannot be left invariant with respect to any structure of homogeneous group), but if we take the first order expansion near* $x = 0$, $e^x - 1 \sim x$, *we get the vector fields*

$$X = \partial_x + 2y\partial_t; Y' = \partial_y - 2x\partial_t$$

which are 1*-homogeneous and left invariant on the Heisenberg group* $\mathbb{H}^1$. *We note that:*

1) $X, Y', [X, Y']$ *are three independent vectors at any point, exactly like* $X, Y, [X, Y]$. *In other words,* Y *and* Y' *are indistinguishable from the point* *of view of the Lie algebra structure, up to the step which is required to check Hörmander's condition.*

2) Moreover, near $x = 0$ we have

$$Y - Y' = -2\left(e^x - 1 - x\right)\partial_t.$$

The interesting property of this vector field is that, with respect to the homogeneities of the Heisenberg group (x has weight 1, t has weight 2), $Y - Y'$ is "approximately homogeneous of order 0 near $x = 0$", since $-2\left(e^x - 1 - x\right) \sim -x^2$ and $-x^2\partial_t$ is homogeneous of degree 0, while Y' is homogeneous of degree 1. This implies that if Γ is, say, a $(2 - Q)$-homogeneous function, $Y'\Gamma$ will be $(1 - Q)$-homogeneous, that is more singular than Γ, while $(Y - Y')\,\Gamma$ will be approximately $(2 - Q)$-homogeneous, that is as singular as Γ. As we will see later in more detail, this is the key point which makes this approximation useful.

Generalizing the above example, the idea is to take the Taylor polynomials of some fixed order, in a neighborhood of some point x_0, of the coefficients of the vector fields X_i and to approximate each X_i with the corresponding Y_i having polynomial coefficients. More precisely, if the original vector fields satisfy Hörmander's condition at step s, this means that not more than $s-1$ derivatives of the coefficients need to be computed when checking this condition, hence the vector fields obtained replacing each coefficient with its Taylor polynomial of degree $s - 1$ will satisfy the same relevant commutator relations. Actually, the Lie algebras generated by the two sets of vector fields will be *undistinguishable up to step s*.

Now, we would like that the polynomial vector fields Y_i were 2-homogeneous and left invariant with respect to some structure of homogeneous group. However we know that, on the one hand, the Lie algebras of the X_i's and the Y_i's have the same structure up to step s, while, on the other hand, the Lie algebra generated by a system of homogeneous left invariant vector field always possesses some special property: for instance, its structure must be the same at any point. But this means that we cannot hope to approximate our original system of vector fields with a "good" one unless our original system *already satisfies some additional algebraic condition* which, in particular, makes its Lie algebra "of constant structure". We could express more precisely this condition saying that for any choice of N vector fields, among $X_1, X_2, ..., X_q$ and their commutators up to step s, if these vectors are independent at *some* point then they must be independent at *any* point.

Now, Rothschild-Stein's idea is *to add extra variables to our original vector fields*, in order to fulfil this condition.

Example 2.2 *Let us consider Grushin's vector fields*

$$X_1 = \partial_x, X_2 = x\partial_y \text{ which "live" in } \mathbb{R}^2$$

and generate a Lie algebra which has not the same structure at any point, because X_1, X_0 are independent if and only if $x \neq 0$. Starting with these vector fields, *we can build the new ones*

$$\widetilde{X}_1 = \partial_x, \widetilde{X}_2 = X_2 + \partial_t = x\partial_y + \partial_t \text{ which "live" in } \mathbb{R}^3.$$

Note that $\widetilde{X}_1, \widetilde{X}_0, \left[\widetilde{X}_1, \widetilde{X}_0\right]$ *are independent at any point of* $\mathbb{R}^3$*. Their Lie algebra is the same as that of the Heisenberg group* $\mathbb{H}^1$*, and actually a smooth change of variables in* $\mathbb{R}^3$ *can turn these vector fields into the "canonical form"* $X' = \partial_{x'} + 2y'\partial_{t'}, Y' = \partial_{y'} - 2x'\partial_{t'}$ *of* $\mathbb{H}^1$*. The vector fields* $\widetilde{X}_1, \widetilde{X}_2$ *satisfy the desired condition, moreover they project onto the original* X_1, X_2*, in the sense that for any function* $f(x, y)$ *which does not depend on the added* t *variable, we have*

$$X_1 f = \widetilde{X}_1 f; X_2 f = \widetilde{X}_2 f.$$

This property should make easy to get the desired a priori estimates for $X_i X_j u$ *once we have proved analogous estimates for* $\widetilde{X}_i \widetilde{X}_j u$ *in a higher dimensional space.*

We say that the vector fields X_1, X_2 have been *lifted* to $\widetilde{X}_1, \widetilde{X}_2$. In the above simple example, the lifted vector fields are already left invariant and homogeneous. More generally one expects to build up a two-step process:

Example 2.3 *Let us consider the operator*

$$L = X_1^2 + X_2^2 \text{ with}$$
$$X_1 = \partial_x, X_2 = (e^x - 1)\,\partial_y \text{ in } \mathbb{R}^2.$$

(The vector fields have the same structure than Grushin's vector fields, but non-polynomial coefficients). Then:

First step: we lift X_1, X_2 *to*

$$\widetilde{X}_1 = \partial_x, \widetilde{X}_2 = (e^x - 1)\,\partial_y + \partial_t \text{ in } \mathbb{R}^3.$$

Note that $\widetilde{X}_1, \widetilde{X}_2, \left[\widetilde{X}_1, \widetilde{X}_2\right]$ *are independent at any point of* $\mathbb{R}^3$*.*

Second step: we approximate $\widetilde{X}_1, \widetilde{X}_2$ *with*

$$Y_1 = \partial_x, Y_2 = x\partial_y + \partial_t$$

which are left invariant and 1*-homogeneous in* $\mathbb{H}^1$ *(up to a smooth change of coordinates).*

2.2 Lifting

It is time to reformulate the previous discussion giving a precise definition and stating a theorem.

Definition 2.4 *We say that a system of smooth vector fields* $Z_1, Z_2, ..., Z_q$ *is* free up to step s *in a domain* Ω *of* $\mathbb{R}^N$ *if the* Z_i*'s and their commutators up to step* s *do not satisfy any linear relation other than those which hold automatically as a consequence of antisymmetry of the Lie bracket and Jacobi identity.*

To put it into another way, $Z_1, Z_2, ..., Z_q$ are free up to step s if and only if the only relations of linear dependence which we can write among them and their commutators up to step s (at any point of Ω), are those which can be established *without knowing the coefficients* of the Z_j's.

If the vector fields satisfy Hörmander's condition at step s and are free up to step s, then in particular their Lie algebra has "constant structure". As we will see with the examples, however, the converse may not be true.

Example 2.5 *1. The "usual" vector fields X, Y on $\mathbb{H}^1$ are free up to step 2:*

$$X, Y, [X, Y]$$

are linearly independent, and there are not other commmutators to be considered, up to step 2.

2. The Grushin vector fields

$$X = \partial_x, Y = x\partial_y$$

are not free at step 2 because

$$Y = x\,[X, Y]$$

which is a nontrivial relation between a generator and a commutator of step 2. In this case, as already noted, the choice of a natural basis of $\mathbb{R}^2$ is different from point to point.

3. The "usual" vector fields X_1, X_2, Y_1, Y_2 on $\mathbb{H}^2$ (see Example 1.9) are not free up to step 2, because

$$[X_1, Y_1] = -4\partial_t = [X_2, Y_2]$$

and this is a nontrivial relation between two commutators of step 2. Note that in this case the Lie algebra actually has *the same structure at any point.*

We can now state the lifting theorem proved by Rothschild-Stein in [44]. To simplify the language, we state it only for operators "sum of squares".

Theorem 2.6 (Lifting, see [44, Thm. 4]) *Let $X_1, ..., X_q$ be vector fields satisfying Hörmander's condition of step s at x_0. (This clearly implies that such property holds in a suitable neighborhood of x_0). Then there exist an integer m and vector fields $\widetilde{X}_k$ defined in a neighborhood of $(x_0, 0) \in \mathbb{R}^{n+m}$, of the form*

$$\widetilde{X}_k = X_k + \sum_{j=1}^{m} u_{kj}\left(x, t_1, t_2, ..., t_{j-1}\right)\partial_{t_j}$$

$(k = 1, ..., q)$, *where the u_{kj}'s are polynomials, such that the $\widetilde{X}_k$'s are free up to step s and satify Hörmander's condition at step s in this neighborhood. (Meaning that the $\widetilde{X}_k$ and their commutators up to step s span $\mathbb{R}^{n+m}$).*

The proof of this theorem is long and very technical, and for the moment we will not say anything about it. Instead, we shall explain the role of this result in the general strategy to get L^p estimates for general Hörmander operators.

2.3 Approximation with left invariant vector fields

Starting from a system of vector fields $X_1, X_2, ..., X_q$ which satisfy Hörmander's condition at step s in a neighborhood of $x_0 \in \mathbb{R}^n$, we have therefore built a new family of "lifted" vector fields $\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_q$, which are *free up to step* s and still satisfy Hörmander's condition at step s in a neighborhood of $(x_0, 0) \in \mathbb{R}^{n+m}$. The second step of the theory consists in approximating these free vector fields with homogeneous left invariant vector fields on a suitable homogeneous group.

To approach this problem, we start with some algebraic remarks. Disregarding the explicit form of the vector fields $\widetilde{X}_k$ in cartesian coordinates, the structure of their Lie algebra up to step s (that is, the number and type of independent objects among the $\widetilde{X}_k$'s and their commutators up to step s) is completely determined by the requirement of being *free up to step* s. This also means that the dimension $N = n + m$ of the lifted space only depends on the numbers q and s.

Example 2.7 *If* $q = 2$ *and* $s = 3$ *we have to consider the following independent vector fields:*

$$\widetilde{X}_1, \widetilde{X}_2; \left[\widetilde{X}_1, \widetilde{X}_2\right]; \left[\widetilde{X}_1, \left[\widetilde{X}_1, \widetilde{X}_2\right]\right]; \left[\widetilde{X}_2, \left[\widetilde{X}_1, \widetilde{X}_2\right]\right].$$

Since these vector fields are 5*, this means that* $N = 5$*. We actually don't know whether the higher order commutators, like*

$$\left[\widetilde{X}_1, \left[\widetilde{X}_1, \left[\widetilde{X}_1, \widetilde{X}_2\right]\right]\right]$$

vanish, or satisfy other nontrivial relations (this depends on the actual explicit form of the vector fields in cartesian coordinates). However, up to step 3, their algebra is determined by the numbers $q = 2$ *and* $s = 3$*.*

Now the idea is that the Lie algebra of homogeneous left invariant vector fields Y_k which (hopefully) approximate locally the $\widetilde{X}_k$'s can be defined abstractly as the *free Lie algebra of step* s *on* q *generators*, which by definition is the quotient of the free Lie algebra on q generators with the ideal spanned by the commutators of length at least $s + 1$. This means that the vector fields Y_k and their commutators up to step s satisfy exactly the same relations than the $\widetilde{X}_k$'s, but moreover all their commutators of step $> s$ vanish. The structure of homogeneous group $\mathbb{G}$ in $\mathbb{R}^N$ is, in turn, determined by that of the corresponding Lie algebra. The vector fields $\widetilde{X}_k$ are defined in a neighborhood of $\xi_0 = (x_0, 0) \in \mathbb{R}^N$; the vector fields Y_k are defined in the whole $\mathbb{R}^N$, and their behavior near the origin will approximate the behavior of the $\widetilde{X}_k$ near ξ_0. This means that the approximation between $\widetilde{X}_k$ and Y_k is realized *in a suitable system of coordinates*. The precise approximation result proved by Rothschild-Stein is the following:

Theorem 2.8 (Approximation, see [44, Thm. 5]) *Assume* $\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_q$*, are free up to step* s *and satisfy Hörmander's condition at step* s *in a neighborhood of* $(x_0, 0) \in \mathbb{R}^{n+m} = \mathbb{R}^N$*. There exist a structure of homogeneous group* $\mathbb{G}$

on $\mathbb{R}^N$, $N = n + m$, a family of homogeneous left invariant Hörmander's vector fields $Y_1, Y_2, ..., Y_q$ on $\mathbb{G}$ and, for any η in a neighborhood of $(x_0, 0)$, a smooth diffeomorphism

$$\xi \mapsto \Theta_\eta(\xi)$$

from a neighborhood of η onto a neighborhood of the origin in $\mathbb{G}$, smoothly depending on η, such that for any smooth function $f : \mathbb{G} \to \mathbb{R}$,

$$\widetilde{X}_i \left(f\left(\Theta_\eta(\cdot)\right)\right)(\xi) = \left(Y_i f + R_i^\eta f\right)\left(\Theta_\eta(\xi)\right) \tag{2.2}$$

($i = 1, ..., q$) where the "remainder" R_i^η is a smooth vector fields of local degree ≤ 0 *and smoothly depends on the parameter η.*

The previous statement needs to be completed by the following

Definition 2.9 *A differential operator P on $\mathbb{G}$ is said to have* local degree less than or equal to k *if, after taking the Taylor expansion at* 0 *of its coefficients, each term obtained is homogeneous of degree $\leq k$.*

Also, the role of $\Theta_\eta(\xi)$ is better explained by the following

Theorem 2.10 *Let*

$$\rho(\xi, \eta) = \left\|\Theta_\eta(\xi)\right\|,$$

where $\|\cdot\|$ is a homogeneous norm on $\mathbb{G}$. Then:

(i) is a quasidistance *(see Definition 1.12);*

(ii) ρ is locally equivalent to the control distance induced by the vector fields $\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_q$.

(iii) The relation between ρ and the Euclidean distance is expressed by

$$c_1 |\xi - \eta| \leq \rho(\xi, \eta) \leq c_2 |\xi - \eta|^{1/s}$$

where s is the step at which Hörmander's condition holds.

(iv) Denoting the ρ-balls with the symbol $B(\xi, R)$ we have

$$c_1 R^Q \leq |B(\xi, R)| \leq c_2 R^Q$$

for R small enough. The same holds for the balls corresponding to the control distance induced by the vector fields $\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_q$. By comparison, recall that for the balls corresponding to the control distance of the vector fields $Y_1, ..., Y_q$ in $\mathbb{G}$ one has exactly

$$|B_Y(u, R)| = cR^Q.$$

Remark 2.11 *The simple behavior of the volume of metric balls $B_{\widetilde{X}}$ in this case heavily depends on the fact that the vector fields $\widetilde{X}_i$ are free. As we will see, the geometry related to a general system of Hörmander's vector fields can be much more involved. Note that in the original paper by Rothschild-Stein the notion of control distance of a general system of vector fields is not present. They just use the function ρ, the Euclidean distance, and the natural distance on the group $\mathbb{G}$, $d(u, v) = \left\|v^{-1} \circ u\right\|$.*

In order to better understand in which sense the vector field R_i^η in (2.2) can be seen as a "small remainder", let us consider the action of $\widetilde{X}_i$ on a function $f\left(\Theta_\eta\left(\xi\right)\right)$ when f is homogeneous of some negative degree $-\alpha$ on $\mathbb{G}$ and smooth ouside the origin, as happens when f is Folland's fundamental solution Γ on $\mathbb{G}$, or a first-order derivative derivative $Y_i\Gamma$. In this case we have

$$\begin{gathered} Y_i f \quad (-\alpha-1)\text{-homogeneous, hence} \\ \left|Y_i f\left(\Theta_\eta\left(\xi\right)\right)\right| \leq \frac{c}{\left\|\Theta_\eta\left(\xi\right)\right\|^{\alpha+1}} = \frac{c}{\rho\left(\xi,\eta\right)^{\alpha+1}}, \end{gathered}$$

while $R_i^\eta f$ is not a homogeneous function, but nevertheless, since R_i^η has local degree ≤ 0,

$$\left|R_i^\eta f\left(\Theta_\eta\left(\xi\right)\right)\right| \leq \frac{c}{\left\|\Theta_\eta\left(\xi\right)\right\|^{\alpha}} = \frac{c}{\rho\left(\xi,\eta\right)^{\alpha}}$$

for ξ near η, that is: *this term is less singular than $Y_i f$.*

The three theorems we have just stated (lifting; approximation; properties of the map Θ), together with some more properties of the map $\Theta_\eta\left(\xi\right)$ proved in [44] and which we will recall when appropriate, constitute a powerful tool, known as "Rothschild-Stein's lifting and approximation technique", which has proved to have further applications than the one for which it was originally designed, and has actually been used by several authors, as we will illustrate in §3. The general idea is that, thanks to this technique, the study of local properties of a general Hörmander operator can sometimes be reduced to the study of a homogeneous left invariant Hörmander operator, of the kind studied by Folland in [27].

2.4 Parametrix and L^p estimates

Let us now illustrate haw this machinery is used by Rothschild-Stein in proving a priori L^p estimates for general Hörmander operators. As usual, we will concentrate on the case of an operator "sum of squares of Hörmander's vector fields" (satisfying Hörmander's condition at step s),

$$L = \sum_{i=1}^{q} X_i^2$$

defined and satisfying these assumptions in a neighborhood of some $x_0 \in \mathbb{R}^n$.

We consider the corresponding lifted operator

$$\widetilde{L} = \sum_{i=1}^{q} \widetilde{X}_i^2$$

where the $\left\{\widetilde{X}_i\right\}_{i=1}^q$ are free up to step s and satisfy Hörmander's condition at step s in some neighborhood $\widetilde{\Omega} = \Omega \times I$ of $(x_0, 0) \in \mathbb{R}^{n+m} = \mathbb{R}^N$ (Thm. 2.6).

Let's note that if we are able to prove L^p estimates for $\widetilde{L}$, that is

$$\|u\|_{W^{2,p}_{\widetilde{X}}(\widetilde{\Omega}')} \leq c\left\{\left\|\widetilde{L}u\right\|_{L^p(\widetilde{\Omega})} + \|u\|_{L^p(\widetilde{\Omega})}\right\} \tag{2.3}$$

for $1 < p < \infty, \widetilde{\Omega}' \Subset \widetilde{\Omega}$, $u \in C^\infty\left(\widetilde{\Omega}\right)$, where

$$\|u\|_{W^{2,p}_{\widetilde{X}}} = \sum_{i,j=1}^{q}\left\|\widetilde{X}_i\widetilde{X}_j u\right\|_{L^p} + \sum_{k=1}^{q}\left\|\widetilde{X}_k u\right\|_{L^p} + \|u\|_{L^p}$$

then we immediately get similar estimates for L, that is the original "unlifted operators". Namely, applying the previous estimate to

$$u(x,t) = f(x)$$

with $f \in C^\infty(\Omega)$ and recalling that $X_i f = \widetilde{X}_i f$, we get

$$\|f\|_{W^{2,p}_{\widetilde{X}}(\Omega')} \leq c\left\{\|Lf\|_{L^p(\Omega)} + \|f\|_{L^p(\Omega)}\right\}.$$

Hence, it is enough to prove (2.3) in the lifted space. Let us consider the structure of homogeneous group $\mathbb{G}$ in $\mathbb{R}^N$, the corresponding 1-homogeneous left invariant vector fields Y_i in $\mathbb{G}$ and the map $\Theta_\eta(\xi)$, defined for ξ, η in a neighborhood of $\xi_0 = (x_0, 0)$, as appear in Thm. 2.8. The operator

$$\mathcal{L} = \sum_{i=1}^{q} Y_i^2$$

is a 2-homogeneous, left invariant, Hörmander operator on $\mathbb{G}$. By Folland's theory [27], it admits a $(2-Q)$-homogeneous, left invariant, fundamental solution Γ smooth outside the pole, such that

$$\mathcal{L}\int_{\mathbb{R}^N}\Gamma\left(v^{-1}\circ u\right) f(v)\,dv = f(u)$$

for every $f \in C_0^\infty\left(\mathbb{R}^N\right)$. Starting with Γ, Rothschild-Stein build a *parametrix* for $\widetilde{L}$. Let us consider the kernel

$$\Gamma\left(\Theta_\eta(\xi)\right)$$

and let us compute

$$\widetilde{L}\left[\Gamma\left(\Theta_\eta(\cdot)\right)\right](\xi).$$

Recall that, by Thm. 2.8,

$$\widetilde{X}_i\left[\Gamma\left(\Theta_\eta(\cdot)\right)\right](\xi) = \left(Y_i\Gamma + R_i^\eta\Gamma\right)\left(\Theta_\eta(\xi)\right),$$

hence

$$\begin{aligned}\widetilde{L}\left[\Gamma\left(\Theta_\eta(\cdot)\right)\right](\xi) &= (\mathcal{L}\Gamma)\left(\Theta_\eta(\xi)\right) + \sum_{i=1}^{q}\left(Y_i R_i^\eta \Gamma + Y_i R_i^\eta \Gamma + R_i^\eta R_i^\eta \Gamma\right)\left(\Theta_\eta(\xi)\right) \\ &= \delta_0\left(\left(\Theta_\eta(\xi)\right)\right) + \left(R^\eta \Gamma\right)\left(\Theta_\eta(\xi)\right)\end{aligned}$$

where δ_0 is the Dirac mass and R^η has local degree ≤ 1, therefore, for every $\xi \neq \eta$, ξ near η, we have

$$\left|\left(R^\eta\Gamma\right)\left(\Theta_\eta(\xi)\right)\right| \leq \frac{c}{\left\|\Theta_\eta(\xi)\right\|^{Q-1}} = \frac{c}{\rho(\xi,\eta)^{Q-1}}$$

that is, $\left(R^\eta\Gamma\right)\left(\Theta_\eta(\xi)\right)$ behaves like a *fractional* (nonsingular, locally integrable) integral kernel. The above computation contains the basic idea. More precisely, we need to define the parametrix as

$$K(\xi,\eta) = a(\xi)\,\Gamma\left(\Theta_\eta(\xi)\right) b(\eta)$$

with a, b cutoff functions, with support small enough so that $K(\xi,\eta)$ is defined on the whole space (recall that $\Theta_\eta(\xi)$ is only locally defined). Then we define the integral operator

$$Pf(\xi) = \int K(\xi,\eta)\, f(\eta)\, d\eta$$

and compute, for any test function f, $\widetilde{L}(Pf)$, finding (by the above computation)

$$\widetilde{L}(Pf)(\xi) = a(\xi)\, f(\xi) + \int K_1(\xi,\eta)\, f(\eta)\, d\eta \tag{2.4}$$

where K_1 is a suitable nonsingular (locally integrable) kernel, satisfying a growth estimate like

$$\left|K_1(\xi,\eta)\right| \leq \frac{c}{\rho(\xi,\eta)^{Q-1}}.$$

Since we would like to have a representation formula of f in terms of $\widetilde{L}f$, modulo a nonsingular integral operator, we now *transpose* the identity (2.4), finding

$$P^*\left(\widetilde{L}^* f\right)(\xi) = a(\xi)\, f(\xi) + \int K_1(\eta,\xi)\, f(\eta)\, d\eta.$$

Recall that $\widetilde{L}^* = \widetilde{L} +$ (lower order terms). In our sketch of Rothschild-Stein's argument we decide to ignore this point and simply write $\widetilde{L}^* = \widetilde{L}$, hence

$$a(\xi)\, f(\xi) = P^*\left(\widetilde{L} f\right)(\xi) - \int K_1(\eta,\xi)\, f(\eta)\, d\eta \tag{2.5}$$

where (exchanging the variables ξ, η in the kernel of P),

$$P^* f(\xi) = \int a(\eta)\,\Gamma\left(\Theta_\xi(\eta)\right) b(\xi)\, f(\eta)\, d\eta.$$

We are ready to take two derivatives $\widetilde{X}_i\widetilde{X}_j$ of both sides of (2.5). For any ξ in a small ball where $a \equiv 1$, we have

$$\begin{aligned}\widetilde{X}_i\widetilde{X}_j f(\xi) &= \widetilde{X}_i\widetilde{X}_j P^*\left(\widetilde{L}f\right)(\xi) - \widetilde{X}_i\widetilde{X}_j \int K_1(\eta,\xi) f(\eta)\, d\eta \\ &\equiv I + II.\end{aligned} \tag{2.6}$$

Let us reflect on the two terms at the right-hand side. The term I is the second derivative of an integral operator whose kernel behaves like the fundamental solution Γ: morally speaking, we should get a singular integral; and this is actually the case: exploiting again the Approximation Theorem 2.8, one can prove

$$\begin{aligned}\widetilde{X}_i\widetilde{X}_j P^*\left(\widetilde{L}f\right)(\xi) = P.V.\int a(\eta)\left(Y_iY_j\Gamma\right)\left(\Theta_\xi(\eta)\right) b(\xi)\, \widetilde{L}f(\eta)\, d\eta + \\ + c_{ij}(\xi)\left(\widetilde{L}f\right)(\xi) + (\text{remainders}),\end{aligned} \tag{2.7}$$

where (remainder) is a fractional integral operator (that is, having a locally integrable kernel) acting on $\widetilde{L}f$.

$$II = \widetilde{X}_i\widetilde{X}_j(S_1 f) = \widetilde{X}_i(S_0 f).$$

where the kernel of S_1 grows like ρ^{1-Q} while that of S_0 grows as ρ^{-Q}: S_0 is a singular integral operator.

The expression $\widetilde{X}_i(S_0 f)$ is an unpleasant one. However, Rothschild-Stein are able to prove that

$$II = \widetilde{X}_j(S_0 f) = \sum_{k=1}^{q} S_0^{(k)}\left(\widetilde{X}_k f\right) + S_0' f \tag{2.8}$$

for suitable singular integral operators $S_0^{(h)}$, S_0' of type 1. Proving this relation of "commutation" between vector fields and integral operators is a heavy part of the theory, involving long and subtle reasonings.

This, together with (2.6) and (2.7) shows that

$$\widetilde{X}_i\widetilde{X}_j f(\xi) = S_0\left(\widetilde{L}f\right)(\xi) + c_{ij}(\xi)\left(\widetilde{L}f\right)(\xi) + \sum_{k=1}^{q} S_0^{(k)}\left(\widetilde{X}_k f\right) + S_0' f$$

Finally, taking L^p norms of both sides we get

$$\left\|\widetilde{X}_i\widetilde{X}_j f\right\|_p \le c\left\{\left\|S_0\left(\widetilde{L}f\right)\right\|_p + \left\|\widetilde{L}f(\xi)\right\|_p + \sum_{k=1}^{q}\left\|S_0^{(k)}\left(\widetilde{X}_k f\right)\right\|_p + \|S_0' f\|_p\right\}.$$

Now, suppose we know that operators of type 0 maps L^p into L^p continuously, then we would conclude

$$\left\|\widetilde{X}_i\widetilde{X}_j f\right\|_p \le c\left\{\left\|\widetilde{L}f\right\|_p + \sum_{k=1}^{q}\left\|\widetilde{X}_k f\right\|_p + \|f\|_p\right\}$$

and, introducing the Sobolev norms $W_{\widetilde{X}}^{k,p}$ (with respect to vector fields),

$$\|f\|_{W_{\widetilde{X}}^{2,p}} \leq c\left\{\left\|\widetilde{L}f\right\|_{L^p} + \|f\|_{W_{\widetilde{X}}^{1,p}} + \|f\|_{L^p}\right\}$$

which is (almost) the desired estimate.

2.5 Singular integral estimates

So we are left to check that operators of type 0 are L^p continuous. Recall that these integral operators do not "live" on a homogeneous group but on a (local) *space of homogeneous type*, however a particularly simple one, whose structure can be seen as a small local perturbation of that of $\mathbb{G}$, via the diffeomorphism $u = \Theta_\xi(\eta)$. Namely, recall that:

$$\rho(\xi, \eta) = \|\Theta_\xi(\eta)\|,$$

with $\|u\|$ a homogeneous norm on $\mathbb{G}$, is a quasidistance, and

$$|B(\xi, r)| \simeq r^Q \text{ for small } r.$$

Moreover, the change of variables

$$u = \Theta_\xi(\eta)$$

within an integral, for η fixed, gives

$$d\xi = c(\eta)(1 + O(\|u\|))\, du$$

with $c(\eta)$ smooth and bounded away from zero.

Just to give an idea of how these properties simplify the study of operators of type 0, let us check the cancellation property for the kernel

$$Y_i Y_j \Gamma(\Theta_\eta(\xi)).$$

Let us compute

$$\int_{R_1<\rho(\xi,\eta)<R_2} Y_i Y_j \Gamma(\Theta_\eta(\xi))\, d\xi =$$

letting $u = \Theta_\xi(\eta)$

$$\begin{aligned}
&= c(\eta)\int_{R_1<\|u\|<R_2} Y_i Y_j \Gamma(u)(1 + O(\|u\|))\, du = \\
&= c(\eta)\int_{R_1<\|u\|<R_2} Y_i Y_j \Gamma(u)\, du + c(\eta)\int_{R_1<\|u\|<R_2} Y_i Y_j \Gamma(u)\, O(\|u\|)\, du = \\
&= A + B.
\end{aligned}$$

Now $A = 0$ by the vanishing property which holds in view of Folland's study of the homogeneous fundamental solution Γ on homogeneous groups, while

$$|Y_i Y_j \Gamma(u)\, O(\|u\|)| \leq \frac{c}{\|u\|^{Q-1}},$$

hence a standard computation shows that

$$|B| \leq c$$

uniformly in R_1, R_2. Hence we also have

$$\left| \int_{R_1 < \rho(\xi,\eta) < R_2} Y_i Y_j \Gamma\left(\Theta_\eta\left(\xi\right)\right) d\xi \right| \leq c$$

for any $R_2 > R_1 > 0$.

The singular integral theory already used by Folland in [27], that is Coifman-Weiss [25] results coupled with those by Knapp-Stein [36], are still adaptable to this situation[1], and L^p continuity of operators of type 0 can be actually proved. This concludes our account of Rothschild-Stein's paper.

2.6 Some final comments on the quest of a-priori estimates in Sobolev spaces

Let us point out some points in the paper by Rothschild-Stein which motivate further research in this field.

1. In contrast with the result proved in the case of homogeneous groups, here a priori estimates are only *local*. Proving L^p estimates for $X_i X_j u$ on the whole $\mathbb{R}^n$ for general Hörmander operators is a difficult problem.

2. A priori estimates are stated, in the paper, with the language of *regularity results*: "if Lu belongs to this space, then u belongs to that space", and not actually writing a priori estimates. If one writes down the results in terms of a priori estimates, one finds something like:

$$\|f\|_{W_X^{2,p}(B_r)} \leq c \left\{ \|Lf\|_{L^p(B_{2r})} + \|f\|_{W_X^{1,p}(B_{2r})} + \|f\|_{L^p(B_{2r})} \right\}.$$

Now, "taking to the left-hand side" the term $\|f\|_{W_X^{1,p}(B_{2r})}$ is not a trivial task with these Sobolev spaces. This issue has been addressed and answered many years later, see for instance the papers [14], [15], [20].

3. So far, the best a priori estimates that we can hope to get on a domain Ω are something like

$$\|f\|_{W_X^{2,p}(\Omega')} \leq c \left\{ \|Lf\|_{L^p(\Omega)} + \|f\|_{L^p(\Omega)} \right\} \text{ for } \Omega' \Subset \Omega,$$

but not for $\Omega' = \Omega$. This poses the question of proving estimates *near the boundary* in Sobolev norms: however, this is a completely open, difficult, problem.

4. No explicit statement is done in [44] about the dependence of the constants. For instance, one could ask: if, for a system $(X_i)_{i=1}^q$ of Hörmander's

[1] We have written that the theory is *adaptable*, not *directly applicable* to our situation. Rothschild-Stein [44] just sketch a proof of this adaptation. By now, however, this point is clearly established on the basis of more recent theories of singular integrals.

vector fields, we consider an operator

$$L_A = \sum_{i,j=1}^{q} a_{ij} X_i X_j$$

where $A = (a_{ij})_{i,j=1}^{q}$ is a constant, symmetric, positive matrix, so that L_A can be actually rewritten as a Hörmander operator, can we say that a-priori estimates hold with some constants depending on the matrix A only through the minimum and maximum eigenvalues? The answer cannot easily be read from the paper [44]. The interested reader can find something more on this topic in [11, §5.3], [13].

Let us also spend some words about the *levels of generality* of the researches about Hörmander operators. The three papers by Folland-Stein [31], Folland [27], Rothschild-Stein [44], besides containing fundamental ideas, results and techniques which are currently used in the research on this field, are also representative of three different levels of generality in the study of Hörmander operators, which still charachterize the current research:

1) sublaplacians on Heisenberg groups;
2) Hörmander operators on homogeneous groups;
3) general Hörmander operators.

Each of these environments has its open problems and challenges. Moreover, there are typical differences in the kind of results which are usually proved in different contexts. For instance, in homogeneous groups one hopes to prove global results, while for general systems of Hörmander's vector fields one usually confines to local results.

There is also a finer scale of generality which appears in some lines of research: from the less to the more general context one can study Hörmander's vector fields on:

- the Heisenberg group $\mathbb{H}^1$;
- Heisenberg groups $\mathbb{H}^n$;
- groups of Heisenberg type (also called H-groups);
- stratified groups of step 2;
- stratified (Carnot) groups;
- homogeneous (graded) Lie groups;
- Lie groups with polynomial growth;
- general Lie groups;
- general Hörmander's vector fields.

The possible requirement that the vector fields be free is a further assumption which can be made both in the context of groups and in the general situation.

Another important difference in generality, which is transversal with respect to the previous gerarchyzation, consists in studying sum of squares of Hörmander's vector fields *or* Hörmander operators containing a drift X_0. The possible presence of the drift X_0 can create deep additional problems. By the way I just mention that Kolmogorov-Fokker-Planck operators, which constitute an important motivation for this theory (see for instance [11, §2.1)]), actually contain the drift.

So far, the literature devoted to general Hörmander's vector fields is considerably narrower than that dealing with Hörmander's vector fields on some kind of Lie group, as the literature devoted to Hörmander operators (with drift) is considerably narrower than that dealing with sum of squares.

3 Further applications of the lifting theorem, and variations on the theme

3.1 First alternative proofs and applications of the lifting theorem (1970s-1980s)

The paper by Rothschild-Stein is deep, long and not easy to read. One reason for this is that it actually contains a number of new deep ideas, expressed by highly technical results. As a consequence, within the proofs many details are left to the reader. The lifting theorem is one of these new deep ideas, and a result of independent interest. In the years immediately following Rothschild-Stein's paper, several authors have given alternative proofs of it. In 1978 Hörmander-Melin [34], and Goodman [32], independently, presented alternative proofs of the lifting theorem and a pointwise version of the approximation theorem.

Folland [28], 1977, aiming to present a more transparent proof of the lifting theorem, actually got a different result. He considered the special case of Hörmander's vector fields defined in the whole $\mathbb{R}^n$, whose Lie algebra is nilpotent and homogeneous, assumptions which are fulfilled for instance by the Grushin operator. In this situation he proved that the vector fields can be lifted to (not necessarily free!), *left invariant* and homogeneous vector fields on a Carnot group, without the necessity of introducing a remainder. In other words, in this special case the relation (2.2) simplifies to

$$\widetilde{X}_i \left(f \left(\Theta_\eta \left(\cdot \right) \right) \right) = \left(Y_i f \right) \left(\Theta_\eta \left(\cdot \right) \right).$$

Moreover, differently from what happens in the case of general Hörmander vector fields, Folland's lifting is *global*. In [28] Folland's motivation was more the simplification of the proof than some new application. As an application, the Author generalizes an example, due to Baouendi-Goulaouic, of hypoelliptic, but not analytical hypoelliptic, operator. However, in very recent years, some new and interesting application of Folland's lifting theorem has been given by Biagi-Bonfiglioli [1], 2017, and [2], in the explicit construction of a global homogeneous fundamental solution for a class of homogeneous (but not translation invariant) Hörmander operators on $\mathbb{R}^n$, therefore generalizing Folland's result in [27]. We will give later some more details about these results in §3.2.4.

3.1.1 The control distance

Before going on, we need to recall the definition of *control distance* induced by the vector fields $X_1, X_2, ..., X_q$ in $\mathbb{R}^N$. This is a fairly general concept, making sense for every system of Lipschitz continuous vector fields in a domain $\Omega \subset \mathbb{R}^N$.

For any $\delta > 0$, let $C_{x,y}(\delta)$ be the class of absolutely continuous mappings $\varphi : [0,1] \longrightarrow \Omega$ which satisfy

$$\varphi'(t) = \sum_{i=1}^{q} a_i(t) (X_i)_{\varphi(t)} \text{ a.e.}$$

with $a_i : [0,1] \to \mathbb{R}$ measurable functions,

$$|a_i(t)| \leqslant \delta \text{ a.e.}$$
$$\varphi(0) = x, \varphi(1) = y.$$

Then define

$$d(x,y) = \inf \{\delta > 0 : \exists \varphi \in C_{x,y}(\delta)\},$$

with the convention $\inf \emptyset = +\infty$. We also define the d-balls

$$B_X(x,r) = \{y \in \Omega : d(x,y) < r\}.$$

Whenever $X_1, X_2, ..., X_q$ is a system of (smooth) Hörmander's vector fields in Ω, then the famous *Chow-Rashevskii connectivity theorem* assures that actually $d(x,y)$ is finite for every couple of points. Roughly speaking, this means that every couple of points in Ω can be joined by a curve composed of integral lines of $X_1, ..., X_q$.

Moreover, for Hörmander's vector fields one can prove that locally the distance d satisfies the following relation with the Euclidean distance:

$$c_1 |x-y| \leq d(x,y) \leq c_2 |x-y|^{1/s}$$

where s is the step of Hörmander's condition.

If $X_1, X_2, ..., X_q$ is the system of generators of a Carnot group, then the control distance d has several extra properties:

i) d is left invariant:

$$d(z \circ x, z \circ y) = d(x,y) \quad \forall x,y,z \in \mathbb{R}^N; \tag{3.1}$$

ii) d is 1-homogeneous:

$$d(D_\lambda(x), D_\lambda(y)) = \lambda d(x,y) \quad \forall x,y \in \mathbb{R}^N, \lambda > 0;$$

iii) since the Lebesgue measure in $\mathbb{R}^N$ is the Haar measure of $\mathbb{G}$, the volume of metric balls satisfies the simple relation

$$|B(x,r)| = |B(0,1)|\ r^Q,$$

for every $x \in \mathbb{G}$ and $r > 0$, where Q is the homogeneous dimension of $\mathbb{G}$;

iv) the function

$$\|x\| = d\,(x,0)$$

is a homogeneous norm. More precisely, it also satisfies the stronger properties

$$\begin{aligned} \left\|x^{-1}\right\| &= \|x\| \\ \|x \circ y\| &\leqslant \|x\| + \|y\|\,. \end{aligned} \tag{3.2}$$

By contrast, we are going to show that for a general system of Hörmander vector fields, the volume of control balls does not behave as a fixed power of the radius.

3.1.2 Geometry of control balls and size estimates on the fundamental solution

The first important application of the lifting theorem, after Rothschild-Stein's paper, is contained in the fundamental paper by Nagel-Stein-Weinger [42], 1985. Let us describe some of the ideas contained in [42].

We begin fixing some notation which will be useful in the following. Let $X_1, X_2, ..., X_q$ be a system of vector fields satisfying Hörmander's condition at step s in some open connected Ω of $\mathbb{R}^n$. For any multiindex $I = (i_1, i_2, ..., i_k)$ of *length* $|I| = k$ we set:

$$X_I = X_{i_1} X_{i_2} ... X_{i_k}$$

and

$$X_{[I]} = \left[X_{i_1}, \left[X_{i_2}, ... \left[X_{i_{k-1}}, X_{i_k}\right] ...\right]\right].$$

If $I = (i_1)$, then

$$X_{[I]} = X_{i_1} = X_I.$$

As usual, $X_{[I]}$ can be seen either as a differential operator or as a vector field. We will write

$$X_{[I]} f$$

to denote the differential operator $X_{[I]}$ acting on a function f, and

$$\left(X_{[I]}\right)_x$$

to denote the vector field $X_{[I]}$ evaluated at the point x. By Hörmander's condition, the vectors

$$\left\{\left(X_{[I]}\right)_x\right\}_{|I| \leq s}$$

span $\mathbb{R}^n$ for any $x \in \Omega$.

The main result contained in [42] is *an estimate of the volume of metric balls*, that is balls with respect to the control distance, induced by the vector fields (see §3.1.1). In turn, this study is intertwined to the study of another notion of distance induced by the vector fields, which we are going to recall.

By Hörmander's condition, the vectors

$$\left\{\left(X_{[I]}\right)_x\right\}_{|I|\leq s}$$

span $\mathbb{R}^n$ for any $x \in \Omega$. Hence for any absolutely continuous curve $\varphi : [0,1] \longrightarrow \Omega$ there exist measurable functions $\{a_I(t)\}_{|I|\leq r}$ defined in $[0,1]$ such that

$$\varphi'(t) = \sum_{|I|\leq s} a_I(t)\left(X_{[I]}\right)_{\varphi(t)} \text{ a.e. } t \in [0,1].$$

With this in mind, we can define the *subelliptic metric* introduced by Nagel-Stein-Wainger [42]:

Definition 3.1 *For any $\delta > 0$, let $C_{x,y}^{(1)}(\delta)$ be the class of absolutely continuous mappings $\varphi : [0,1] \longrightarrow \Omega$ which satisfy*

$$\varphi'(t) = \sum_{|I|\leq s} a_I(t)\left(X_{[I]}\right)_{\varphi(t)} \quad a.e.$$

with $a_I : [0,1] \to \mathbb{R}$ measurable functions,

$$|a_I(t)| \leq \delta^{|I|},$$
$$\varphi(0) = x, \varphi(1) = y.$$

Then define

$$d_1(x,y) = \inf\left\{\delta > 0 : \exists\varphi \in C_{x,y}^{(1)}(\delta) \text{ with } \varphi(0) = x, \varphi(1) = y\right\}.$$

Differently from the control distance d introduced in §3.1.1, the *finiteness* of $d_1(x,y)$ for every couple of $x, y \in \Omega$ (Ω open connected subset of $\mathbb{R}^n$) is now obvious. Also d_1 is a distance.

The idea behind the above definition is the following: to reach a point y starting from x we can follow *any* curve we want, but we move faster if we choose to follow integral lines of the basic vector fields $X_1, ..., X_q$ ("the highways") and we move slower and slower as we follow integral lines of commutators of the X_i's of higher and higher step ("minor roads"). The distance d_1 measures the total *time* we spend to reach y from x. By comparison, the control distance d is defined using only "the highways" to connect points, and in that case the connectivity property is not obvious. One of the deep results proved by Nagel-Stein-Weinger in [42] is that:

Theorem 3.2 *The distances d and d_1 are locally equivalent:*

$$d_1(x,y) \leq d(x,y) \leq c d_1(x,y)$$

locally.

Inequality $d_1(x,y) \leq d(x,y)$ is trivial, because any $C_{x,y}(\delta) \subset C^{(1)}_{x,y}(\delta)$. Inequality $d(x,y) \leq cd_1(x,y)$ in particular implies the finiteness of $d(x,y)$ (since $d_1(x,y)$ is obviously finite), that is the connectivity theorem. In some sense, this equivalence can be seen as a quantitative version of the connectivity theorem.

In particular, in order to estimate the volume of the metric balls $B_d(x,r)$ we can equivalently estimate the volumes of the balls $B_{d_1}(x,r)$, which turns out to be easier.

In order to study the volume of metric balls and other metric properties, it is useful to exploit suitable coordinate systems.

First, let us recall that the *exponential of a vector field* is defined, as usual, as follows: we say that

$$\exp(X)(x_0) = f(1)$$

where $f(t)$ is the solution to the Cauchy problem

$$\begin{cases} f'(t) = X_{f(t)} \\ f(0) = x_0 \end{cases}$$

Now, let $\left\{ \left(X_{[I]}\right)_{x_0} \right\}_{I \in \mathcal{B}}$ be a basis of $\mathbb{R}^n$ obtained choosing, at the point x_0, suitable commutators of the X_i's. Let us consider the map

$$\{u_I\}_{I\in\mathcal{B}} \mapsto x = \exp\left(\sum_{I\in\mathcal{B}} u_I X_{[I]} \right)(x_0) \tag{3.3}$$

defined from a neighborhood of the origin in the $\mathbb{R}^n$ space of the variables $\{u_I\}_{I\in\mathcal{B}}$ to a neighborhood $U(x_0)$.

The jacobian of this map at $u = 0$ equals the matrix $\left\{ \left(X_{[I]}\right)_{x_0} \right\}_{I \in \mathcal{B}}$, which is nonsingular since the vectors $\left(X_{[I]}\right)_{x_0}$ are a basis of $\mathbb{R}^n$; hence this map is a local diffeomorphism, which represents a neighborhood of x_0 by means of the canonical coordinates $\{u_I\}_{I\in\mathcal{B}}$. This coordinate system depends on the choice $\mathcal{B}$ of the basis. Note that, by definition of the distance d_1,

$$|u_I| \leq \delta^{|I|} \ \forall I \in \mathcal{B} \Longrightarrow x \in B_{d_1}(x_0, \delta).$$

Hence, to study d_1-metric balls, an interesting object is the δ-box

$$\left\{ u \in \mathbb{R}^n : |u_I| \leq \delta^{|I|} \ \forall I \in \mathcal{B} \right\}$$

(which also depends on our choice of $\mathcal{B}$). Now, under the exponential mapping (3.3) the volume of the image of this box should be comparable to

$$\delta^{|\mathcal{B}|} \left| \det \left\{ \left(X_{[I]}\right)_{x_0} \right\}_{I\in\mathcal{B}} \right| \tag{3.4}$$

having set

$$|\mathcal{B}| = \sum_{I\in\mathcal{B}} |I|.$$

This suggests that, in order to find a sharp estimate of the volume of metric balls $B_{d_1}(x_0, \delta)$, one has to choose the set $\mathcal{B}$ of generators that maximizes the quantity (3.4). Note that this choice depends *both on the center and on the radius of the ball*, which makes very delicate the analysis carried out by Nagel-Stein-Wainger. Their result is the following:

Theorem 3.3 (Volume of metric balls) *For any $\Omega' \Subset \Omega$ there exist constants C_1, C_2, r_0 such that for any $x \in \Omega'$ and $\delta \leq r_0$ one has*

$$0 < C_1 \leq \frac{|B_{d_1}(x, \delta)|}{\Lambda(x, \delta)} \leq C_2$$

where

$$\Lambda(x, \delta) = \sum_{\mathcal{B}} \left| \det \left\{ \left(X_{[I]} \right)_x \right\}_{I \in \mathcal{B}} \right| \delta^{|\mathcal{B}|}$$

and the sum is taken over all the possible n-tuples $\mathcal{B}$ of multiindices I with $|I| \leq r$.

As already remarked, the function $|B_{d_1}(x, \delta)|$ is locally equivalent to $|B_d(x, \delta)|$. Henceforth we will simply write $|B(x, \delta)|$ to denote one of these two functions.

Example 3.4 *Let us consider the Grushin vector fields:*

$$X_1 = \partial_x; X_2 = x\partial_y \text{ in } \mathbb{R}^2.$$

There are just two interesting choices of $\mathcal{B}$:

$\mathcal{B}_1 = ((1), (2))$ *(i.e. X_1, X_2) which gives $|\mathcal{B}_1| = 2$ and* $\left| \det \left\{ \left(X_{[I]} \right) \right\}_{I \in \mathcal{B}_1} \right| = |x|$

$\mathcal{B}_2 = ((1), (1, 2))$ *(i.e. X_1 and $[X_1, X_2]$) which gives $|\mathcal{B}_2| = 3$ and* $\left| \det \left\{ \left(X_{[I]} \right) \right\}_{I \in \mathcal{B}_2} \right| = 1.$

The first choice is possible only at points $x \neq 0$. Hence, at any point $(0, y_0)$ we will have to choose $\mathcal{B}_2$, getting

$$\delta^{|\mathcal{B}_2|} \left| \det \left\{ \left(X_{[I]} \right)_{(0, y_0)} \right\}_{I \in \mathcal{B}_2} \right| = \delta^3$$

while as soon as we move to (x_0, y_0) with $x_0 \neq 0$, choosing $\mathcal{B}_1$ (which has smaller length than $\mathcal{B}_2$) we get

$$\delta^{|\mathcal{B}_1|} \left| \det \left\{ \left(X_{[I]} \right)_{(x_0, y_0)} \right\}_{I \in \mathcal{B}_1} \right| = \delta^2 |x_0|.$$

The above theorem states that

$$C_1 \left(\delta^3 + \delta^2 |x| \right) \leq |B((x, y), \delta)| \leq C_2 \left(\delta^3 + \delta^2 |x| \right)$$

with C_1, C_2 depending on an upper bound on δ and $\sqrt{x^2 + y^2}$. In particular, the balls of center $(0, y_0)$ have volume comparable to δ^3, while the balls of center (x_0, y_0) with large x_0 and small radius δ have volume comparable to δ^2.

A relevant consequence of the above theorem is the following:

Corollary 3.5 (Local doubling property) *For any $\Omega' \Subset \Omega$ there exist constants C, r_0 such that for any $x \in \Omega'$ and $\delta \leq r_0$ one has*

$$|B(x, 2\delta)| \leq C |B(x, \delta)|.$$

Note that, strictly speaking, this does not allow to conclude that (Ω, d, dx) is a space of homogeneous type in the sense of Coifman-Weiss (see Definition 1.12). Actually, the above doubling condition holds only *locally* in Ω.

It is worthwhile to note what happens in the particular case of a system of vector fields which are free up to step s and satisfy Hörmander's condition at step s in Ω. In this case, if $\mathcal{B}$ is a choice of generators such that

$$\det \left\{ \left(X_{[I]} \right)_x \right\}_{I \in \mathcal{B}} \neq 0 \tag{3.5}$$

at some point $x \in \Omega$, then this will be true at *every* point of Ω. Moreover, for a fixed $\Omega' \Subset \Omega$ the function $\left| \det \left\{ \left(X_{[I]} \right)_x \right\}_{I \in \mathcal{B}} \right|$ will have positive lower and upper bounds, hence the function $\Lambda(x, \delta)$ will be equivalent to $c\delta^Q$ for some Q which is the smallest value of $|\mathcal{B}|$ such that (3.5) holds. Therefore:

Corollary 3.6 (Volume of metric balls for free vector fields) *If the X_i's are free up to step s and satisfy Hörmander's condition at step s in Ω then there exists a positive integer Q and, for any $\Omega' \Subset \Omega$, there exist positive constants C_1, C_2, r_0 such that for any $x \in \Omega'$ and $\delta \leq r_0$ one has*

$$C_1 \delta^Q \leq |B(x, \delta)| \leq C_2 \delta^Q.$$

The last part of the paper [42] applies the previous results to the context studied by Rothschild-Stein in [44]. Let $X_1, X_2, ..., X_q$ be any system of Hormander's vector fields in some neighborhood U of $x_0 \in \mathbb{R}^n$; we can consider the distance $d(x, y)$ induced in U by the X_i's; we will call it d_X now, to distinguish from other distances. Let $\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_q$ be the free lifted vector fields defined as in [44] in a domain $\widetilde{U} = U \times I \subset \mathbb{R}^{n+m} = \mathbb{R}^N$. We can consider the distance $d_{\widetilde{X}}$ defined by the vector fields $\widetilde{X}_i$ in $\widetilde{U}$. Also, we know that in $\widetilde{U}$ there is a quasidistance $\rho(\xi, \eta) = \|\Theta_\eta(\xi)\|$ naturally attached to the $\widetilde{X}_i$'s, so two questions arise: which relation does exist between d_X and $d_{\widetilde{X}}$ and between ρ and $d_{\widetilde{X}}$? Nagel-Stein-Wainger prove that:

Proposition 3.7 *The distance $d_{\widetilde{X}}$ and the quasidistance ρ, both defined in subsets of $\mathbb{R}^N$, are locally equivalent.*

More subtle is the relation between $d_{\widetilde{X}}$ and d_X, which are defined in subsets of spaces of different dimensions. First, they prove that

Proposition 3.8 *The projection of the metric ball $B_{\widetilde{X}}((x, t), r)$ on $\mathbb{R}^n$ is $B_X(x, r)$, which means that*

$$d_{\widetilde{X}}((x, t), (y, s)) \geq d_X(x, y).$$

It is not possible, however, to prove a control in the reverse sense, at the level of distances. For instance, a "reasonable" inequality like

$$d_{\widetilde{X}}\left((x,0),(y,0)\right) \leq c d_X(x,y)$$

has never been proved. What one can prove is a control (in both senses) at the level of *volumes*. This is a deep result, relying on the previous analysis of the structure of metric balls, and reads as follows:

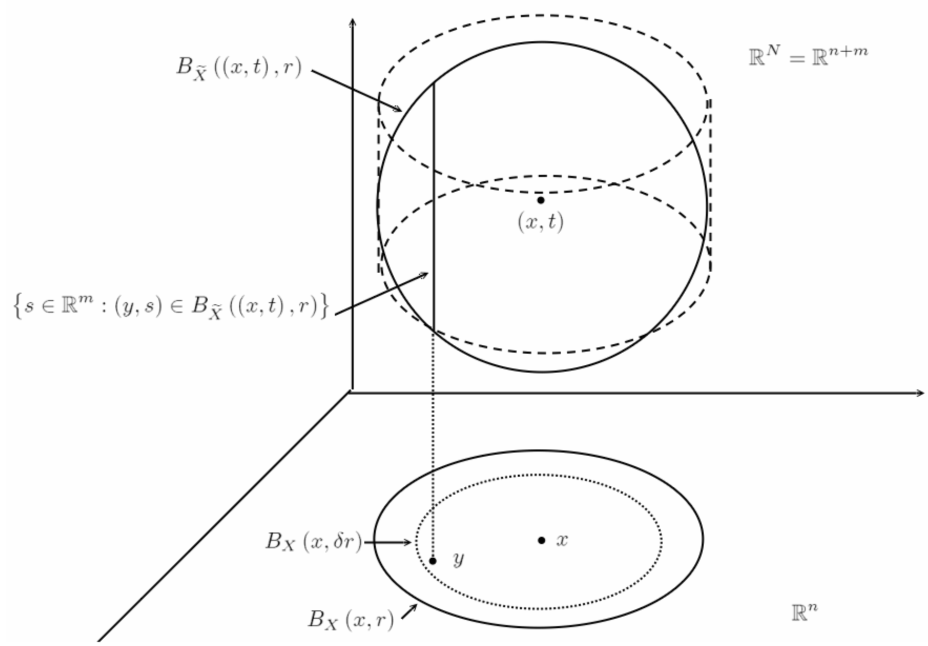
Theorem 3.9 (Volumes of lifted and unlifted balls) *There exists $c_1 > 0$ such that for any $r > 0$ small enough and any $(x,t) \in \mathbb{R}^N, y \in \mathbb{R}^n$ (in the small neighborhoods under consideration),*

$$\left|B_{\widetilde{X}}\left((x,t),r\right)\right| \geq c_1 \left|B_X(x,r)\right| \left|\left\{s \in \mathbb{R}^m : (y,s) \in B_{\widetilde{X}}\left((x,t),r\right)\right\}\right|.$$

Conversely, there exists $\delta \in (0,1)$ and $c_2 > 0$ such that, for any r, (x,t) as above and $y \in B_X(x,\delta r)$,

$$\left|B_{\widetilde{X}}\left((x,t),r\right)\right| \leq c_2 \left|B_X(x,r)\right| \left|\left\{s \in \mathbb{R}^m : (y,s) \in B_{\widetilde{X}}\left((x,t),r\right)\right\}\right|.$$

In the above statements, the symbol $|\cdot|$ stands for the full dimensional Lebesgue measure in any of the three spaces $\mathbb{R}^N, \mathbb{R}^n, \mathbb{R}^m$.



The geometric meaning of the theorem is that the volume of the lifted ball $B_{\widetilde{X}}$ is equivalent to the volume of a "cylinder" having the ball B_X as basis and the set $\left\{s \in \mathbb{R}^m : (y,s) \in B_{\widetilde{X}}\right\}$ as "height". However, this equivalence of volumes is not the consequence of a simple set inclusion, but more the substitute of such lacking inclusion.

The above theorem is a powerful tool, which helps in deriving estimates in the original space starting from analogous estimates in the lifted space, where they are more easily established.

The final result proved by Nagel-Stein-Wainger is a direct application of Theorem 3.9. Assume we have a kernel $k(\xi,\eta)$ in the lifted space, satisfying the local bound

$$k((x,t),(y,s)) \leq c \frac{d_{\widetilde{X}}((x,t),(y,s))^a}{\left|B_{\widetilde{X}}\left((x,t), d_{\widetilde{X}}((x,t),(y,s))\right)\right|}$$

and let us define the "restricted kernel", obtained from the previous one saturating by integration the lifted variables:

$$Rk(x,y) = \int_{\mathbb{R}^m} k((x,0),(y,s))\,\phi(s)\,ds$$

for a suitable cuoff function ϕ. Then it is not difficult to prove, using Theorem 3.9 and the doubling condition, that

$$Rk(x,y) \leq c \frac{d_X(x,y)^a}{\left|B_{\widetilde{X}}(x, d_X(x,y))\right|}.$$

Also, if

$$\left|\widetilde{X}_{i_1}\widetilde{X}_{i_2}...\widetilde{X}_{i_h} k((x,t),(y,s))\right| \leq c \frac{d_{\widetilde{X}}((x,t),(y,s))^{a-h}}{\left|B_{\widetilde{X}}\left((x,t), d_{\widetilde{X}}((x,t),(y,s))\right)\right|}$$

then

$$\left|X_{i_1}X_{i_2}...X_{i_h} Rk(x,y)\right| \leq c \frac{d_X(x,y)^{a-h}}{\left|B_{\widetilde{X}}(x, d_X(x,y))\right|}.$$

The previous general fact applies in particular to the parametrix built in [44] for the lifted operator, satisfying

$$\Gamma(\Theta((x,t),(y,s))) \leq \frac{c}{\rho((x,t),(y,s))^{Q-2}} = c \frac{d_{\widetilde{X}}((x,t),(y,s))^2}{\left|B_{\widetilde{X}}\left((x,t), d_{\widetilde{X}}((x,t),(y,s))\right)\right|}.$$

Then the function

$$\gamma(x,y) = \int_{\mathbb{R}^m} \Gamma(\Theta((x,0),(y,s)))\,\phi(s)\,ds$$

satisfies a bound

$$\gamma(x,y) \leq c \frac{d_X(x,y)^2}{\left|B_{\widetilde{X}}(x, d_X(x,y))\right|},$$

and similar bounds hold for the derivatives:

$$\left|X_{i_1}X_{i_2}...X_{i_h}\gamma\left(x,y\right)\right| \leq c\frac{d_X\left(x,y\right)^{2-h}}{\left|B_{\widetilde{X}}\left(x,d_X\left(x,y\right)\right)\right|}$$

near the pole. This is what the Authors actually prove. What they *claim* is that these bounds apply, near the pole, to the fundamental solution of the original Hörmander operator $\sum X_i^2$. This statement is a bit unsatisfactory, in my opinion. Actually, in [42] it is implicitly understood that we can apply this procedure to a fundamental solution for the lifted operator, obtaining a fundamental solution γ for the original operator. However, in [44] no fundamental solution is exhibited for the lifted operator, but only the parametrix $\Gamma\left(\Theta\left(\left(x,t\right),\left(y,s\right)\right)\right)$, which satisfies a simple bound by the homogeneity of Γ. Saturating this parametrix should produce a parametrix for the original Hörmander operator, and only with a hard work one could produce, starting with this, an effective fundamental solution. If, instead, one starts with an effective fundamental solution γ of the original Hörmander operator, whose existence is assured by some abstract argument, then one should prove that this γ behaves near the pole like the kernel that we get by saturation.

Apart from the above minor criticism, the idea of exploiting Thm 3.9 to derive bounds on some kernel obtained by saturating the lifted variables of another kernel defined in a higher dimensional space, is deep, and has been fruitfully used also by other authors. We will give some account about this in §§3.2.3-3.2.4.

Let us mention that, independently and almost simultaneously to [42], Sanchez-Calle published the paper [45], 1984, with the same final goal (local estimates on the fundamental solution) and also containing theorem 3.9, which is proved by a completely different method, relying on previous results by Fefferman-Phong. Sanchez-Calle's paper does not contain the general study of the volume of metric balls performed by Nagel-Stein-Weinger. On the other hand, differently from the paper by Nagel-Stein-Weinger, it contains a construction of the fundamental solution which is then shown to satisfy suitable bounds.

Let us also mention the paper by Fefferman, Sánchez-Calle, 1986 [26], which contains a far-reaching extension of the aforementioned upper bounds on fundamental solutions, for general subelliptic operators with nonnegative characteristic form (not necessarily written as sum of squares of vector fields).

3.1.3 Jerison's Poincaré inequality for Hörmander's vector fields

A second important application of the lifting theorem is contained in the famous paper by Jerison [35], 1986, where the following *Poincaré's inequality for Hörmander's vector fields* is proved: if $X_1,...,X_q$ is a system of Hörmander's vector fields in a domain Ω, and $B\left(x,r\right)$ denote the balls with respect to the control distance induced by the vector fields, then there exist constant C,r_0 such that for every $x \in \Omega$ and $r < r_0$ such that $B(x, 2r) \subset \Omega$ we have

$$\min_{a \in \mathbb{R}} \int_{B(x,r)} |f(y) - a|^2 \, dy \leq Cr^2 \int_{B(x,r)} \sum_{i=1}^{q} |X_i f(y)|^2 \, dy$$

for every $f \in C^\infty \left(\overline{B(x,r)} \right)$. This result relies both on Rothschild-Stein's lifting and approximation theorem, and on Nagel-Stein-Weinger results (in particular, Thm 3.9). Moreover, in order to describe precisely the dependence of the constants C, r_0, Jerison scrutinizes the proofs of these two results, to state quantitative versions of them.

Let us give some ideas of the strategy of Jerison's proof. First, the Author establishes Poincaré's inequality on Carnot groups. This actually is not too difficult. Second, the Author makes the following statement. Assume that we have proved the desired Poincaré's inequality under the additional assumption that the vector fields are free. Then, in particular, this inequality holds for the lifted vector fields $\widetilde{X}_1, \widetilde{X}_2, ..., \widetilde{X}_q$, in the sense of Rothschild-Stein, that is

$$\min_{a \in \mathbb{R}} \int_{\widetilde{B}(\xi,r)} |f(\eta) - a|^2 \, d\eta \leq Cr^2 \int_{\widetilde{B}(\xi,r)} \sum_{i=1}^{q} \left| \widetilde{X}_i f(\eta) \right|^2 d\eta,$$

with the obvious meaning of symbols. Applying this inequality to any smooth function $f(y, s)$ which actually does not depend on s we find

$$\min_{a \in \mathbb{R}} \int_{\widetilde{B}(\xi,r)} |f(y) - a|^2 \, dyds \leq Cr^2 \int_{\widetilde{B}(\xi,r)} \sum_{i=1}^{q} |X_i f(y)|^2 \, dyds$$

which, thanks to Theorem 3.9, easily gives

$$\min_{a \in \mathbb{R}} \int_{B(x,r)} |f(y) - a|^2 \, dy \leq Cr^2 \int_{B(x,r)} \sum_{i=1}^{q} |X_i f(y)|^2 \, dy.$$

So, thanks to Nagel-Stein-Weinger's result, the proof is reduced to showing that Poincaré's inequality holds for free Hörmander's vector fields, already knowing that it holds on Carnot groups. This can be done applying Rothschild-Stein's lifting and approximation result, but not easily. Actually, by a clever application of these techniques Jerison manages to prove the following "main Lemma":

$$\min_{a \in \mathbb{R}} \int_{\widetilde{B}(\xi,r/c)} |f(\eta) - a|^2 \, d\eta \leq Cr^2 \left(\int_{\widetilde{B}(\xi,cr)} \sum_{i=1}^{q} \left| \widetilde{X}_i f(\eta) \right|^2 d\eta + \int_{\widetilde{B}(\xi,r)} |f(\eta)|^2 \, d\eta \right),$$

for some constant $c > 1$. In other words, what one gets is a relation similar to Poincaré's inequality, but with an "error term" $\int_{\widetilde{B}(\xi,r)} |f(\eta)|^2 \, d\eta$ apparently bigger than the left-hand side. The redeeming feature of this inequality, however, is the (possibly small) r^2 which multiplies this error term. Thanks to a convering theorem of Whitney type, then, the error term is eventually adsorbed and the desired result is proved.

3.2 More recent variations on the theme and applications of the lifting theorem (2000s-today)

Here we briefly discuss some more recent extensions or further applications of the lifting theorem, together with some background and motivation.

3.2.1 A weighted version of the lifting theorem

In 1999 Christ-Nagel-Stein-Wainger in [24, § 22] proved a more general version of the lifting theorem, because they also consider "weighted" vector fields. Assume that $X_1, ..., X_q$ is a system of vector fields in an open set $\Omega \subseteq \mathbb{R}^n$, and each X_i has an assigned "weight" expressed by an integer $p_i \geq 1$. We say that a commutator $[X_i, X_j]$ has weight $p_i + p_j$, and so on. Assume that $X_1, ..., X_q$ satisfy Hörmander's condition at weighted step s, that is the commutators of weighted length $\leq s$ are enough to span $\mathbb{R}^n$ at every point of Ω. Then there exists an integer $N = n + m$ and lifted vector fields $\widetilde{X}_1, ..., \widetilde{X}_q$,

$$\widetilde{X}_i = X_i + \sum_{j=1}^{m} b_{ij}\left(x, t_1, ..., t_{j-1}\right) \partial_{t_j}$$

($x \in \Omega$) which are free up to weighted step s and still statisfy Hörmander's condition at weighted step s. The Authors also prove an approximation result. This more general theorem in particular covers the case of a Hörmander operator with drift,

$$\sum_{i=1}^{q} X_i^2 + X_0$$

where $X_1, ..., X_1$ have weight 1 while X_0 has weight 2. (This case is not explicitly carried out in Rothschild-Stein's paper, although it is required by their theory: this is one of the details left to the reader in [42]). In [24, § 22] this lifting theorem is one of the tools used to prove the L^p boundedness of singular Radon transforms and their maximal analogues. These operators are kinds of singular integral operators, which however involve integration over a k-dimensional submanifold of $\mathbb{R}^n$, varying from point to point, with $k < n$. So, this more general lifting result is proved for a specific reason, not just for the seak of generality. We will not get into more details about this.

3.2.2 Another version of the lifting theorem for groups

In 2005 Bonfiglioli-Uguzzoni in [9] proved the following lifting theorem for Carnot groups. Let $X_1, ..., X_q$ be the generators of a Carnot group $\mathbb{G}$ in $\mathbb{R}^n$, which is not free. (The Heisenberg groups $\mathbb{H}^n$ for $n \geq 2$ are an example of non-free Carnot groups). Then there exists a higher dimensional free Carnot group $\mathbb{G}'$ in $\mathbb{R}^N$ for some $N > n$ such that the generators $\widetilde{X}_1, ..., \widetilde{X}_q$ of $\mathbb{G}'$ are a lifting of $X_1, ..., X_q$. Like for Folland's 1977 result in [27], here vector fields are lifted directly to the generators of a Carnot group, without the necessity of a remainder. Hence, this result is not covered by the original one by Rothschild-Stein. Differently from Folland, here the lifted vector fields are free, and the starting vector fields are already left invariant. So, this result and Folland's one are not comparable. Let us explain the Authors' motivation for this result.

In the series of papers [6], 2002, [7], 2004, [10], 2007, by Bonfiglioli-Lanconelli-Uguzzoni, the Authors have deeply investigated the family of operators

$$\mathcal{H} = \sum_{i,j=1}^{q} a_{ij}(x,t) X_i X_j - \partial_t \tag{3.6}$$

where $X_1, ..., X_q$ are the generators of a Carnot group on $\mathbb{R}^n$, the matrix $\{a_{ij}\}$ is uniformly positive on a domain, and the coefficients a_{ij} are Hölder continuous. The aim is to build a fundamental solution (heat kernel) of this variable coefficient operators, and prove that it satisfies sharp Gaussian bounds. This construction is performed in [7] and, exploiting this heat kernel, in [10] a scale-invariant Harnack inequality is proved. The heat kernel for H is built by the parametrix method, starting with the heat kernel of a corresponding constant coefficient operator

$$H_A = \sum_{i,j=1}^{q} a_{ij} X_i X_j - \partial_t. \tag{3.7}$$

In turn, to make this construction possible, it is necessary to prove that the heat kernel of the constant coefficient operator H_A satisfies uniform Gaussian estimates, as the constant matrix $A = \{a_{ij}\}$ ranges in the class $\mathcal{A}_\nu$ of matrices satisfying

$$\nu |\xi|^2 \leq \sum_{i,j=1}^{q} a_{ij} \xi_i \xi_j \leq \nu^{-1} |\xi|^2$$

for some $\nu \in (0,1)$, every $\xi \in \mathbb{R}^q$. This is performed in [6]. It is also necessary to show that the heat kernel depends Hölder-continuously on the matrix a_{ij}. Proving this poses a subtle problem, which the Authors overcome in [6], also exploiting the results in [8], 2004, as follows. The Authors construct a Lie-group diffeomorphism which transforms the heat operator

$$h = \sum_{i=1}^{q} X_i^2 - \partial_t$$

into the operator H_A in (3.7). The heat kernel Γ_A of H_A can be expressed by means of this diffeomorphism and the matrix A in terms of the (fixed) heat kernel γ of h. This explicit relation allows to prove the desired continuous dependence of Γ_A on A. But the construction of this diffeomorphism r*equires the extra assumption that the Carnot group be free.* If the Carnot group is not free, the Authors apply the lifting theorem in [9] to make the construction possible. The general case is then reduced to the free one.

We end this survey of this line of research pointing out that all the results about the fundamental solution and Harnack inequality for operators $\mathcal{H}$

as in (3.6) have been extended to the case of general Hörmander's vector fields (without any underlying group structure) by Bramanti-Brandolini-Lanconelli-Uguzzoni in [16], 2010.

3.2.3 Nonsmooth Hörmander's vector fields

In 2010 Bramanti-Brandolini-Pedroni in [18] extended the lifting and approximation theorem to the case of *nonsmooth Hörmander vector fields.* This result has its context within a recent new theory which is worthwhile to be briefly explained here. If

$$X_i = \sum_{i,j=1}^{n} b_{ij}\left(x\right) \partial_{x_j}$$

(for $i = 1, 2, ..., q$) is a system of vector fields defined in a domain $\Omega \subseteq \mathbb{R}^n$, in order to compute their commutators up to some length r it is enough that the coefficients b_{ij} be $(r-1)$-times continuously differentiable. So it is meaningful to consider a set of vector fields with coefficients $C^{s-1}(\Omega)$ for some positive integer s, satisfying Hormander's condition at step s in Ω. We will say that this is a system of nonsmooth (actually, C^{s-1}) Hörmander vector fields. This is the setting considered in the papers by Bramanti-Brandolini-Pedroni [18], 2010, [19], 2013, and in the paper [17], by Bramanti-Brandolini-Manfredini-Pedroni, 2017. Actually, to obtain some of the results, this minimal regularity assumption (C^{s-1}) has to be reinforced to $C^{s-1,\alpha}$ for some $\alpha \in (0,1)$ or to $C^{s-1,1}$ (depending on the result).

In [19] a Poincaré's inequality is proved for a system of $C^{s-1,1}$ vector fields (satisfying Hörmander's condition at step s), thus extending Jerison's result in [35] to the nonsmooth context. (Let us note that a similar result has been obtained, independently and with a different technique, by Morbidelli-Montanari in [40]). Like in Jerison's paper, the result is first established for free (but, in this case, *nonsmooth*) vector fields, then projected on the space of the original variables, thanks to a suitable nonsmooth version of Nagel-Stein-Weinger's theorem 3.9, about the comparison of volumes of lifted and unlifted balls. The required lifting theorem is the one established in [18], under the mere C^{s-1} assumption, and is an adaptation to the nonsmooth case of the proof of the (smooth) lifting theorem given by Hörmander-Melin in [34]. Poincaré's inequality for nonsmooth free vector fields is obtained with a strategy which is different from the one used by Jerison, and instead exploits the techniques contained in the paper by Lanconelli-Morbidelli [39], 2000.

Let us spend a few words also about the *approximation theorem* which is combined with the nonsmooth lifting theorem in [18]. If one tried to apply the same technique used by Rothschild-Stein directly to the nonsmooth lifted vector fields, the resulting map $\Theta_\eta(\xi)$ which plays the role of local coordinates in which $\widetilde{X}_i$ is approximated by the left invariant homogeneous Y_i would possess a poor regularity, and would not serve the scope. This forces us to use a two-step approximation procedure. Assume the vector fields X_i have regularity $C^{s-1,\alpha}$ for some $\alpha \in (0,1)$. Their lifted vector fields $\widetilde{X}_i$ will have the same regularity. Now, at a fixed point $\overline{\xi}$, take the Taylor expansion of order $(s-1)$ of the coefficients of X_i, and consider the smooth vector fields $S_i^{\overline{\xi}}$ having as coefficients these Taylor polynomials. One can prove that the smooth vector fields $S_1^{\overline{\xi}},...,S_q^{\overline{\xi}}$ are still free and satisfy Hörmander's condition at step s. By Rotschild-Stein's approximation theorem, they are locally well approximated by the generators $Y_1,...,Y_q$ of a Carnot group in the lifted space. Now the $\widetilde{X}_i$ are locally approximated by the $S_i^{\overline{\xi}}$ which in turn are locally approximated by the Y_i. The map $\Theta_\eta(\xi)$ which connects $\widetilde{X}_i$ with Y_i now has asymmetric properties with respect to the variables ξ,η: for η fixed, it is a smooth diffeomorphism in ξ; but it depends on η only in C^α-way.

Besides its role in the proof of Poincaré's inequality for nonsmooth vector fields, carried out in [19], this nonsmooth lifting and approximation machinery has been exploited in [17], in a way that is similar to how Nagel-Stein-Weinger have employed it in [42]. In [17] the Authors studied the nonsmooth Hörmander operator

$$L=\sum_{i=1}^{q}X_i^2+X_0.$$

For expositive reasons here we restrict ourselves to the special case

$$L=\sum_{i=1}^{q}X_i^2.$$

Again, it is assumed that for some integer $s\geq 2$ and some $\alpha\in(0,1]$, the vector fields X_i have $C^{s-1,\alpha}(\Omega)$ coefficients and satisfy in $\Omega\subseteq\mathbb{R}^n$ Hörmander's condition at step s. By the nonsmooth lifting and approximation result, we can build a sublaplacian $\sum_{i=1}^{q}Y_i^2$ on a Carnot group in the space $\mathbb{R}^N=\mathbb{R}^{n+m}$ of lifted variables, which by Folland's results possesses a global homogeneous fundamental solution Γ. Let $\Theta_\eta(\xi)$ be the nonsmooth version of Rothschild-Stein map, as built in [18], and let us define the following kernel, by saturation of the lifted variables:

$$P(x,y)=\int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^m}\Gamma\left(\Theta_{(y,k)}(x,h)\right)\phi(h)\,dh\right)\phi(k)\,dk,$$

where ϕ is a suitable cutoff function. Here

$$\xi=(x,h),\eta=(y,k)\in\mathbb{R}^N=\mathbb{R}^{n+m}.$$

This $P(x,y)$ is smooth in x and only Hölder continuous in y. It turns out that P is a parametrix for L and, starting with P, the classical Levi's parametrix method allows to build a local fundamental solution γ for L, with γ and its first order derivatives (with respect to x) satisfying natural growth estimates. Requiring one more degree of regularity in the coefficients of X_i, that is $C^{s,\alpha}(\Omega)$, one can prove that γ also has second order derivatives $X_iX_j\gamma$ satisfying natural growth estimates, and a local solvability result for L is proved. This paper is the first one where general nonsmooth Hörmander operators are studied, and is a good example of pushing forward the techniques of Rothschild-Stein and Nagel-Stein-Weinger.

3.2.4 Homogeneous Hörmander operators

In 2017 Biagi-Bonfiglioli, in [1], considered the class of sum of squares of Hörmander's vector fields $X_1, ..., X_p$ which are defined on the whole $\mathbb{R}^n$ and are homogeneous with respect to some family $\{\delta_\lambda\}_{\lambda>0}$ of dilations (but not necessarily left invariant with respect to a group of translations). The simplest example of this kind is the Grushin operator

$$L = \partial^2_{xx} + x^2 \partial^2_{yy}$$

in $\mathbb{R}^2$. Other examples are the following:

L	$\mathbb{R}^n$	δ_λ	q
$(\partial_{x_1})^2 + (x_1^k\, \partial_{x_2})^2$	2	$(\lambda x_1, \lambda^{k+1} x_2)$	$k+2$
$(\partial_{x_1})^2 + (x_1 \partial_{x_2} + x_2 \partial_{x_3} + \ldots + x_{n-1}\partial_{x_n})^2$	n	$(\lambda x_1, \lambda^2 x_2, \cdots, \lambda^n x_n)$	$\frac{n(n+1)}{2}$
$(\partial_{x_1})^2 + \left(x_1\, \partial_{x_2} + x_1^2\, \partial_{x_3}\right)^2$	3	$(\lambda x_1, \lambda^2 x_2, \lambda^3 x_3)$	6

By Folland's lifting theorem in [28], suitably adapted by the Authors, the operator L can be lifted to the sublaplacian $\widetilde{L}$ of a higher dimensional Carnot group $\mathbb{G}$, which possesses a global, left invariant, homogeneous fundamental solution $\Gamma_{\mathbb{G}}(x, k; y, h) = \Gamma_{\mathbb{G}}\left((x,k)^{-1} * (y,h)\right)$ (here $x, y \in \mathbb{R}^n$, $(x,k), (y,h) \in \mathbb{G}$, $k, h \in \mathbb{R}^m$ are the variables which are added in the lifting procedure). Then the Authors prove that the function

$$\Gamma(x, y) = \int_{\mathbb{R}^m} \Gamma_{\mathbb{G}}\left((x,0)^{-1} \circ (y,h)\right) dh$$

is a well defined, global fundamental solution for the original operator L, enjoying several interesting properties: it is smooth out of the diagonal, symmetric in x, y, strictly positive, locally integrable on $\mathbb{R}^n \times \mathbb{R}^n$; it vanishes when x or y go to infinity; it is jointly homogeneous of degree $2 - q < 0$, i.e.

$$\Gamma(\delta_\lambda(x), \delta_\lambda(y)) = \lambda^{2-q}\Gamma(x, y)$$

for every $\lambda > 0, x \neq y$, where q is the homogeous dimension of $\mathbb{R}^n$, that is

$$q = \sum_{i=1}^n \sigma_i \text{ with } \delta_\lambda(x) = (\lambda^{\sigma_1} x_1, ..., \lambda^{\sigma_n} x_n).$$

This paper represents a very interesting application of Folland's version of the lifting theorem, proved 40 years before. This result has been extended by the same Authors to homogeneous heat-type operators

$$\sum_{i=1}^p X_i^2 - \partial_t$$

in [2]. Note that, in absence of a group structure, very few *global* results are known for Hörmander operators, so these works represent a promising line of research.

Starting with this integral representation of a global fundamental solution for L, it is possible, in turn, to apply Nagel-Stein-Weinger technique (which in this situation can be easily extended to a *global* result) and derive *global* natural growth estimates for Γ. This has been done by Biagi-Bonfiglioli-Bramanti in [3], getting the following global bounds:

$$\left|Z_1 ... Z_r \Gamma(x, y)\right| \leq C \frac{d(x, y)^{2-r}}{\left|B_X(x, d(x, y))\right|}$$

where $Z_1, ..., Z_r$ stand for any r derivatives among $X_1, ..., X_q$, with respect to x *or* y. Also, if $n > 2$

$$C_1 \frac{d(x, y)^2}{\left|B_X(x, d(x, y))\right|} \leq \Gamma(x, y) \leq C_2 \frac{d(x, y)^2}{\left|B_X(x, d(x, y))\right|}$$

(when $n = 2$ the result assumes a more involved form, due to the possibility of logarithmic growth). These results are an interesting example of application of the combination of techniques dating to Rothschild-Stein and Nagel-Stein-Weinger. In the reacher context of homogeneous operators, however, the results are definetely more quantitative than the oroginal ones: here the fundamental solution Γ is a uniquely defined kernel, the constants depend on a few known parameters, the estimates hold globally in $\mathbb{R}^n$.

Analogously, Biagi-Bramanti in [4], derive from the integral representation of the heat kernel of a homogeneous operator $\sum_{i=1}^{q} X_i^2 - \partial_t$, natural global Gaussian estimates.

4 Sketch of the proof of the lifting and approximation theorem

In this last section we will sketch the ideas of the proof of the lifting theorem following the approach followed by Hörmander-Melin [34], 1978; a detailed exposition of the arguments of [34] has been presented in the paper [18], where this approach is adapted to the context of nonsmooth Hörmander vector fields, and will appear in [13].

Once again, here we restrict ourselves to the case of generators of weight one (no drift term). We start recalling once more some notation which has been already used.

Let $X_1, ..., X_q$ be a system of real smooth vector fields, defined in a domain $\Omega \subseteq \mathbb{R}^n$. For any multiindex

$$I = (i_1, i_2, ..., i_k)$$

we define the *weight* (or length) of I as $|I| = k$. For $I = (i_1, i_2, ..., i_k)$ we set:

$$X_I = X_{i_1} X_{i_2} ... X_{i_k}$$

and

$$X_{[I]} = \left[X_{i_1}, \left[X_{i_2}, ... \left[X_{i_{k-1}}, X_{i_k}\right] ...\right]\right].$$

If $I = (i_1)$, then

$$X_{[I]} = X_{i_1} = X_I.$$

As usual, we will write

$$X_{[I]} f$$

to denote the differential operator $X_{[I]}$ acting on a function f, and

$$\left(X_{[I]}\right)_x$$

to denote the vector field $X_{[I]}$ evaluated at the point x.

4.1 Lifting

We are now going to define the concept of *free system of vector fields.* Actually, this notion has been already introduced in §2.2; the definition that we will give in this section is formulated in a more technical way, which actually turns out to be more useful in the proofs. It can be proved that the two definitions are equivalent.

Let us start with the following remark. Any vector field $X_{[I]}$ with $|I| \leq s$ can be rewritten explicitly as a linear combination of operators of the kind X_J for $|J| = |I|$:

$$X_{[I]} = \sum_J A_{IJ} X_J$$

where $[A_{IJ}]_{|I|,|J| \leq s}$ is a matrix of universal constants, built exploiting only those relations between $X_{[I]}$ and X_J which hold automatically, as a consequence of the definition of $X_{[I]}$, regardless of the specific properties of the vector fields $X_1, ..., X_q$. In particular, we see that

$$A_{IJ} = 0 \text{ if } |J| \neq |I| \tag{4.1}$$

and

$$A_{IJ} = \delta_{IJ} \text{ if } |J| = |I| = 1.$$

Example 4.1 *For the system* $\{X_1, X_2\}$ *and* $s = 2$ *we have* 6 *possible multi-indices:*

$$1, 2, (1,1)\ (1,2), (2,1), (2,2).$$

The only nonzero elements of the matrix $\{A_{IJ}\}_{|I|,|J| \leq 2}$ *are:*

$$\begin{aligned} A_{1,1} &= A_{2,2} = 1 \\ A_{(1,2),(1,2)} &= 1 = A_{(2,1),(2,1)} \\ A_{(1,2),(2,1)} &= -1 = A_{(2,1),(1,2)}. \end{aligned}$$

Actually,

$$X_{[(1,2)]} = [X_1, X_2] = X_1 X_2 - X_2 X_1 = X_{(1,2)} - X_{(2,1)}.$$

Note that if $\{a_I\}_{I\in\mathcal{B}}$ is any finite set of constants such that

$$\sum_{I\in\mathcal{B}} a_I A_{IJ} = 0 \ \forall J, \quad \text{then} \quad \sum_{I\in\mathcal{B}} a_I X_{[I]} \equiv 0 \tag{4.2}$$

for arbitrary vector fields $X_1, ..., X_q$, since in this case

$$\sum_{I\in\mathcal{B}} a_I X_{[I]} = \sum_{I\in\mathcal{B}} a_I \sum_J A_{IJ} X_J = \sum_J \left(\sum_{I\in\mathcal{B}} a_I A_{IJ} \right) X_J = 0.$$

Reversing this property we get an alternative definition of free vector fields:

Definition 4.2 (Free vector fields) *For any positive integer σ, we say that the vector fields $X_1, ..., X_q$ are* free up to step σ *at x, if, for any family of constants $\{a_I\}_{|I|\leq\sigma}$,*

$$\sum_{|I|\leq\sigma} a_I \left(X_{[I]}\right)_x = 0 \Longrightarrow \sum_{|I|\leq\sigma} a_I A_{IJ} = 0 \ \forall J.$$

This definition is consistent with the one previously given.

Example 4.3 *(a) In $\mathbb{R}^3$,*

$$\begin{aligned} X_1 &= \partial_{x_1} + 2x_2 \partial_{x_3} \\ X_2 &= \partial_{x_2} - 2x_1 \partial_{x_3} \end{aligned}$$

are free up to step 2 at 0 (Actually, they are free at any point, but for simplicity we check this fact at the origin). Namely, if

$$0 = a_1 (X_1)_0 + a_2 (X_2)_0 + a_{12} ([X_1, X_2])_0 = a_1 \partial_{x_1} + a_2 \partial_{x_2} + a_{12} \partial_{x_3}$$

then $a_1 = a_2 = a_{12} = 0$.

(b) Instead, in $\mathbb{R}^5$,

$$\begin{aligned} X_1 &= \partial_{x_1} + 2x_2 \partial_{x_5} \\ X_2 &= \partial_{x_2} - 2x_1 \partial_{x_5} \\ X_3 &= \partial_{x_3} + 2x_4 \partial_{x_5} \\ X_4 &= \partial_{x_4} - 2x_3 \partial_{x_5} \end{aligned}$$

are not free up to step 2 at 0. Namely,

$$[X_1, X_2] - [X_3, X_4] = 0,$$

which should imply

$$A_{(1,2),J} - A_{(3,4),J} = 0 \ \forall J,$$

while for $J = (1, 2)$ we have

$$A_{(1,2),(1,2)} - A_{(3,4),(1,2)} = 1 - 0 = 1.$$

(c) In $\mathbb{R}^3$,

$$X_1 = x_3\partial_{x_1} + 2x_2\partial_{x_3}$$
$$X_2 = \partial_{x_2} - 2x_1\partial_{x_3}$$

are free up to step 2 *at* $(0,0,1)$*, for in this case the situation is like in example (a), but they are not free up to step* 2 *at* $(0,0,0)$*, since*

$$(X_1)_0 = 0 \text{ but } A_{11} = 1.$$

We also note the following basic facts:

Proposition 4.4 *If the vector fields* $X_1, ..., X_q$ *are free up to step* σ *at some point* x_0*, then there exists a neighborhood* $U(x_0)$ *such that they are free up to step* σ *at any point* $x \in U(x_0)$*.*

Proposition 4.5 *Let* $X_1, ..., X_q$ *be vector fields defined in a neighborhood of* $x \in \mathbb{R}^n$*. If* $\left\{ \left(X_{[I]} \right)_x \right\}_{|I| \leqslant s}$ *span* $\mathbb{R}^n$*, then*

$$n \leqslant \operatorname{rank} [A_{IJ}]_{|I|,|J| \leqslant s}.$$

Moreover, the number $\operatorname{rank} [A_{IJ}]_{|I|,|J| \leqslant s}$ *is, in any case, an absolute constant only depending on the numbers* q, s*.*

The last assertion holds since the matrix $[A_{IJ}]_{|I|,|J| \leqslant s}$ only exploits those relations between $X_{[I]}$ and X_J that hold automatically as a consequence of the definition of $X_{[I]}$, hence it only depends on the step s and the number of vector fields q.

The next proposition contains a deeper property of free vector fields:

Proposition 4.6 *Let* $X_1, ..., X$ *be vector fields defined in a neighborhood of the origin in* $\mathbb{R}^n$ *that are free up to the step* σ *at* 0*. Then for any family of constants* $\{c_I\}_{|I| \leq \sigma} \subset \mathbb{R}$ *there exists a polynomial* u *in* $\mathbb{R}^n$ *such that* $X_I u(0) = c_I$ *when* $|I| \leq \sigma$*.*

The above proposition is perhaps the most technical point of Hörmander-Melin's proof of the lifting theorem. Its proof is delicate, and we will just give an idea of the problem behind it, later.

The next result, instead, contains the main inductive step towards the proof of the lifting theorem, and it is worthwhile to present its proof in detail.

Proposition 4.7 *Let* $X_1, ..., X_q$ *be free of step* $\sigma - 1$ *but not of step* σ *at* 0*. Then one can find vector fields* $\widetilde{X}_j$ *in* $\mathbb{R}^{n+1}$ *of the form*

$$\widetilde{X}_j = X_j + u_j(x) \frac{\partial}{\partial t} \quad (j = 1, ..., q) \tag{4.3}$$

with u_j *polynomial such that*

1. the vector fields $\widetilde{X}_j$ remain free up to the step $\sigma - 1$;
2. for every $s \geq \sigma$,

$$\dim \left\langle \left(\widetilde{X}_{[I]} \right)_0 \right\rangle_{|I| \leq s} = \dim \left\langle \left(X_{[I]} \right)_0 \right\rangle_{|I| \leq s} + 1$$

where the symbol $\langle Y_\alpha \rangle_{\alpha \in \mathcal{B}}$ denotes the vector space spanned by the vectors $\{Y_\alpha : \alpha \in \mathcal{B}\}$.

Before proving the above proposition, let us show how the lifting theorem easily follows from it.

Theorem 4.8 (Lifting) *Let $X_1, ..., X_q$ be vector fields defined in a neighborhood of the origin in $\mathbb{R}^n$ satisfying Hörmander's condition of step s at $x = 0$. Then there exist an integer m and vector fields $\widetilde{X}_k$ in $\mathbb{R}^{n+m}$, of the form*

$$\widetilde{X}_k = X_k + \sum_{j=1}^{m} u_{kj} \left(x, t_1, t_2, ..., t_{j-1} \right) \frac{\partial}{\partial t_j}$$

$(k = 1, ..., q)$, *where the u_{kj}'s are polynomials, such that the $\widetilde{X}_k$'s are free of step s and $\left\{ \left(\widetilde{X}_{[I]} \right)_0 \right\}_{|I| \leq s}$ span $\mathbb{R}^{n+m}$.*

This theorem has an obvious reformulation at any point $x_0 \in \mathbb{R}^n$, with the lifted vector fields defined in a neighborhood of $(x_0, 0) \in \mathbb{R}^{n+m}$. Moreover, in view of Proposition 4.4, both the conclusions of the theorem (freeness and Hörmander's condition at step s for the lifted vector fields) will hold in a suitable neighborhood of this point.

Proof of the lifting theorem. Let $\left\{ \left(X_{[I]} \right)_0 \right\}_{I \in \mathcal{B}}$ be a basis of $\mathbb{R}^n$, for some set $\mathcal{B}$ of n multiindices of weight $\leq s$. Recall that by Proposition 4.5 we have

$$n \leq \operatorname{rank} \left[A_{IJ} \right]_{|I|, |J| \leq s} \equiv c \left(s, q \right), \tag{4.4}$$

an absolute constant only depending on s, q.

Now, let $\sigma \leq s$ be such that $X_1, ..., X_q$ are free of step $\sigma - 1$ but not of step σ, at 0. (If the vector fields X_i were already free of step s, there would be nothing to prove. We also agree to say that the vector fields X_i are free of step 0 if they are not free of step 1). We can then apply Proposition 4.7 and build vector fields

$$\widetilde{X}_j = X_j + u_j \left(x \right) \frac{\partial}{\partial t} \;\; (j = 1, ..., q)$$

in $\mathbb{R}^{n+1}$, free of step $\sigma - 1$ and such that

$$\dim \left\langle \left(\widetilde{X}_{[I]} \right)_0 \right\rangle_{|I| \leq s} = \dim \left\langle \left(X_{[I]} \right)_0 \right\rangle_{|I| \leq s} + 1 = n + 1$$

(because by assumption the $\left\{ \left(X_{[I]} \right)_0 \right\}_{|I| \leq s}$ span $\mathbb{R}^n$). Hence the $\left\{ \left(\widetilde{X}_{[I]} \right)_0 \right\}_{|I| \leq s}$ still span the whole space $\mathbb{R}^{n+1}$. Now: either the vector fields $\left\{ \left(\widetilde{X}_{[I]} \right)_0 \right\}_{|I| \leq s}$

are free of step s, and we are done, or the assumptions of Proposition 4.7 are still satisfied, and we can iterate our argument; in this case, by (4.4), condition $n+1 \leq c(s,q)$ must hold. After a suitable finite number m of iterations, condition $n+m \leq c(s,q)$ cannot hold anymore, and this means that the vector fields $\widetilde{X}_j$ must be free of step s. The iterative argument also shows that the u_{kj}'s are polynomials only depending on the variables $x, t_1, t_2, ..., t_{j-1}$. ■

Proof of Proposition 4.7. Let us show that condition 1 in the above statement holds for any choice of the functions $u_j(x)$ in (4.3). To see this, we first claim that (4.3) implies

$$\widetilde{X}_{[I]} = X_{[I]} + u_I(x) \frac{\partial}{\partial t} \tag{4.5}$$

for any multiindex I and some $u_I \in C^\infty(\mathbb{R}^n)$. Namely, we can proceed by induction on $|I|$. For $|I| = 1$, this is just (4.3); assume (4.5) holds for $|I| = j-1$. For $|I| = j$, let $I = (i, J)$ for some $i = 1, ..., q$ and $|J| = j-1$. Then, by inductive assumption,

$$\begin{aligned}
\widetilde{X}_{[I]} = \widetilde{X}_{[i,J]} = \left[\widetilde{X}_i, \widetilde{X}_{[J]}\right] &= \\
&= \left[X_i + u_i(x) \frac{\partial}{\partial t}, X_{[J]} + u_J(x) \frac{\partial}{\partial t}\right] = \\
&= \left[X_i, X_{[J]}\right] + \left(X_i u_J - X_{[J]} u_i\right) \frac{\partial}{\partial t} = \\
&= X_{[I]} + u_I(x) \frac{\partial}{\partial t}.
\end{aligned}$$

Next, we show that (4.5) implies that the vector fields $\widetilde{X}_i$ are free of step $\sigma - 1$ at 0. If

$$\sum_{|I| \leq \sigma - 1} a_I \left(\widetilde{X}_{[I]}\right)_0 = 0$$

for suitable constants a_I, then by (4.5) we have

$$\begin{aligned}
0 &= \sum_{|I| \leq \sigma-1} a_I \left(X_{[I]} + u_I(x) \frac{\partial}{\partial t}\right)_0 = \\
&= \sum_{|I| \leq \sigma-1} a_I \left(X_{[I]}\right)_0 + \left(\sum_{|I| \leq \sigma-1} a_I u_I(0)\right) \frac{\partial}{\partial t}.
\end{aligned}$$

Since $\frac{\partial}{\partial t}$ is independent from the vectors $\left(X_{[I]}\right)_0$, this implies that

$$\sum_{|I| \leq \sigma-1} a_I u_I(0) = 0$$

and

$$\sum_{|I| \leq \sigma-1} a_I \left(X_{[I]}\right)_0 = 0.$$

But the vector fields X_i are free of step $\sigma - 1$ at 0, hence

$$\sum_{|I| \leq \sigma - 1} a_I A_{IJ} = 0 \text{ for any } J \text{ with } |J| \leq \sigma - 1.$$

Therefore also the vector fields $\widetilde{X}_i$ are free of step $\sigma - 1$ at 0.

We now show that it is possible to choose polynomial functions u_j in (4.3) such that condition 2 in the statement of this proposition holds. To show this, we will prove that there exist polynomials u_j and constants $\{a_I\}_{|I| \leq \sigma}$ such that:

$$\sum_{|I| \leq \sigma} a_I \left(X_{[I]}\right)_0 = 0 \tag{4.6}$$

and

$$\sum_{|I| \leq \sigma} a_I \left(\widetilde{X}_{[I]}\right)_0 \neq 0 \tag{4.7}$$

From (4.6)-(4.7), condition 2 will follow; namely,

$$0 \neq \sum_{|I| \leq \sigma} a_I \left(\widetilde{X}_{[I]}\right)_0 = \sum_{|I| \leq \sigma} a_I \left(\left(X_{[I]}\right)_0 + u_I(0) \frac{\partial}{\partial t}\right) = \left(\sum_{|I| \leq \sigma} a_I u_I(0)\right) \frac{\partial}{\partial t} = b \frac{\partial}{\partial t}$$

with $b \neq 0$, hence

$$\frac{\partial}{\partial t} = \sum_{|I| \leq \sigma} \frac{a_I}{b} \left(\widetilde{X}_{[I]}\right)_0$$

and this shows that

$$\left\langle \left(\widetilde{X}_{[I]}\right)_0 \right\rangle_{|I| \leq s} = \left\langle \left(X_{[I]}\right)_0 \right\rangle_{|I| \leq s} \oplus \left\langle \frac{\partial}{\partial t} \right\rangle,$$

which implies condition 2.

To prove (4.6)-(4.7), we use our assumption on the vector fields X_i: since they are *not* free of step σ, there exist coefficients $\{a_I\}_{|I| \leq \sigma}$ such that (4.6) holds but

$$\sum_{|I| \leq \sigma} a_I A_{IJ} \neq 0 \text{ for some } J \text{ with } |J| \leq \sigma. \tag{4.8}$$

It remains to prove that there exist polynomials u_j such that (4.7) holds for these u_j's and a_I's. To determine these u_j's, let us examine the action of the vector field

$$\sum_{|I| \leq \sigma} a_I \widetilde{X}_{[I]} = \sum_{|I| \leq \sigma} a_I \sum_{|J| \leq \sigma} A_{IJ} \widetilde{X}_J$$

on the function t. For any J with $|J| \leq \sigma$, let us write $J = (J'j)$ for some $j = 1, ..., q$. Then

$$\widetilde{X}_J t = \widetilde{X}_{J'} \widetilde{X}_j t = \widetilde{X}_{J'} \left[\left(X_j + u_j \frac{\partial}{\partial t}\right) t\right] = \widetilde{X}_{J'} u_j = X_{J'} u_j$$

since u_j does not depend on t. We then have:

$$\left(\sum_{|I|\leq\sigma} a_I \widetilde{X}_{[I]}(t)\right)(0) = \sum_{|I|\leq\sigma} a_I \sum_{|J|\leq\sigma} A_{IJ}\left(X_{J'}u_j\right)(0).$$

where in the inner summation, $J = (J'j)$, with $|J'| \leq \sigma - 1$. Since the vector fields X_i are free of step $\sigma - 1$ at 0, by Proposition 4.6 for any choice of constants $\{c_{J'}\}_{|J'|\leq\sigma-1}$ there exists a polynomial $u \in C^{\infty}(\mathbb{R}^n)$ such that $(X_{J'}u)(0) = c_{J'}$. On the other hand, by (4.8), there exists a set of constants $\{c_J\}_{|J|\leq\sigma}$ such that

$$\sum_{|I|\leq\sigma}\sum_{|J|\leq\sigma} a_I A_{IJ} c_J \neq 0.$$

Setting $c_{J'}^{j} = c_J$ if $J = (J'j)$ and applying q times Proposition 4.6 to the q sets of constants $\left\{c_{J'}^{j}\right\}_{|J'|\leq\sigma-1}$, $j = 1, 2, ..., q$, we find polynomials $u_1, ..., u_q$ such that

$$\left(\sum_{|I|\leq\sigma} a_I \widetilde{X}_{[I]}(t)\right)(0) = \sum_{|I|\leq\sigma}\sum_{|J|\leq\sigma} a_I A_{IJ} c_J \neq 0.$$

Hence (4.7) holds. This completes the proof of the proposition. ■

Let us now come back to Proposition 4.6. We want to give the reader an informal idea of the problem involved in its proof, by means of a concrete example. The vector fields of the Heisenberg group $\mathbb{H}^1$ are enough to appreciate the problem. Let

$$\begin{aligned} X_1 &= \partial_{x_1} + 2x_2\partial_{x_3} \\ X_2 &= \partial_{x_2} - 2x_1\partial_{x_3}. \end{aligned}$$

Then the system we have to solve is:

$$\left\{\begin{array}{l} X_1 u(0) = c_1 \\ X_2 u(0) = c_2 \\ X_1^2 u(0) = c_{11} \\ X_1 X_2 u(0) = c_{12} \\ X_2 X_1 u(0) = c_{21} \\ X_2^2 u(0) = c_{22} \end{array}\right. \quad \text{that is} \quad \left\{\begin{array}{l} \partial_{x_1} u(0) = c_1 \\ \partial_{x_2} u(0) = c_2 \\ \partial^2_{x_1 x_1} u(0) = c_{11} \\ \partial^2_{x_1 x_2} u(0) - 2\partial_{x_3} u(0) = c_{12} \\ \partial^2_{x_1 x_2} u(0) + 2\partial_{x_3} u(0) = c_{21} \\ \partial^2_{x_2 x_2} u(0) = c_{22} \end{array}\right. \tag{4.9}$$

Now, a naïf idea to solve this system (or, to better say, to prove in general the solvability of such a system) could be to decouple the system into subgroups of equations, looking for a function u of the kind $u = u_1 + u_2$ with u_i homogeneous polynomial of degree i (in the usual sense), u_1 solving the system of the first two equations and u_2 solving the system of the last four equations. Doing so we would find (due to the vanishing of $\partial_{x_3} u_2(0)$) the generally incompatible conditions

$$\begin{aligned} \partial^2_{x_1 x_2} u_2(0) &= c_{12} \\ \partial^2_{x_1 x_2} u_2(0) &= c_{21}. \end{aligned}$$

We can see that the system (4.9) *is* actually solvable, but in order to solve it we need to choose properly also the value $\partial_{x_3} u(0)$ which is not determined by the group of the first two equations. A solution u is the polynomial:

$$u(x_1, x_2, x_3) = c_1 x_1 + c_2 x_2 + \left(\frac{c_{21} - c_{12}}{4}\right) x_3 + \frac{1}{2}\left[c_{11} x_1^2 + (c_{12} + c_{21}) x_1 x_2 + c_{22} x_2^2\right].$$

The reader can appreciate that a *symmetry issue* arises here: cartesian mixed derivative commute while mixed derivatives with respect to the vector fields do not; this seems to threaten the solvabilty of the system.

A second issue is the following. Trying to prove, in the abstract context, the solvability of the system, we need to exploit the only assumption we have, namely the fact that the vector fields are free of step σ. However, this assumption is formulated in terms of the commutators $X_{[I]}$, while the system itself is written in terms of the differential monomials X_I. Hence in order to exploit our assumption we have to reformulate the problem in a way involving commutators. This will be done seeing both the X_I's and the $X_{[I]}$'s as particular *polynomials* in the vector fields. Actually, the proof of Proposition 4.6 exploits the language of *polynomials in noncommuting variables*, and a key step in the proof will consist in proving the *symmetry* of a suitable j-linear form. We will not go into further details.

4.2 Approximation

Here we want to describe the second part of the procedure related to the lifting theorem, that is the *approximation of free vector fields by left invariant vector fields on a homogeneous group*. By the lifting theorem (Theorem 4.8), starting from any system of vector fields satisfying Hörmander's condition at step s in some neighborhood of the origin in $\mathbb{R}^n$ we can define new vector fields $\widetilde{X}_1, \ldots, \widetilde{X}_q$ in a neighborhood U of $0 \in \mathbb{R}^N \equiv \mathbb{R}^{n+m}$ that are free up to step s at any point of U and such that $\left\{\left(\widetilde{X}_{[I]}\right)_\xi\right\}_{|I| \leq s}$ spans $\mathbb{R}^N$ for any $\xi \in U$. We start with the following

Remark 4.9 *If $\widetilde{X}_1, ..., \widetilde{X}_q$ is a system of vector fields in a bounded domain $U \subset \mathbb{R}^N$, free up to the step s and satisfying Hörmander's condition of step s in U, then it is possible to choose a set $\mathcal{B}$ of N multiindices I with $|I| \leq s$, such that*

$$\left\{\widetilde{X}_{[I]}\right\}_{I \in \mathcal{B}}$$

is a basis of $\mathbb{R}^N$ at every point $\xi \in U$.

We assume this set $\mathcal{B}$ fixed once and for all.

For any $\overline{\xi} \in U$, let us introduce in a neighborhood of $\overline{\xi}$ the set of local ("canonical") coordinates

$$\Phi_{\overline{\xi},\mathcal{B}} : \mathcal{U}(0) \subset \mathbb{R}^N \to \mathbb{R}^N \tag{4.10}$$
$$\Phi_{\overline{\xi},\mathcal{B}}(u) = \exp\left(\sum_{I\in\mathcal{B}} u_I \widetilde{X}_{[I]}\right)\left(\overline{\xi}\right),$$

defined for u in a suitable neighborhood $\mathcal{U}(0)$ of 0. Note that since card $(\mathcal{B}) = N$ we can represent points of $\mathbb{R}^N$ as $(u_I)_{I\in\mathcal{B}}$.

Recall that

$$\frac{\partial \Phi_{\overline{\xi},\mathcal{B}}}{\partial u_J}(0) = \frac{d}{du_J}\left(\exp\left(u_J \widetilde{X}_{[J]}\right)\left(\overline{\xi}\right)\right)_{/u_J=0} = \left(\widetilde{X}_{[J]}\right)_{\overline{\xi}},$$

hence the Jacobian of the map $\Phi_{\overline{\xi},\mathcal{B}}$ at $u = 0$, equals the matrix of the vector fields $\left\{\left(\widetilde{X}_{[I]}\right)_{\overline{\xi}}\right\}_{I\in\mathcal{B}}$ and therefore it is nonsingular.

This allows to define *canonical coordinates* in a suitable neighborhood $U\left(\overline{\xi}\right)$ of $\overline{\xi}$.

Moreover, since the basis $\left\{\left(\widetilde{X}_{[I]}\right)_{\overline{\xi}}\right\}_{I\in\mathcal{B}}$ depends continuously on the point $\overline{\xi}$, the radius of this neighborhood can be taken uniformly bounded away from zero for $\overline{\xi}$ ranging in a compact set.

Henceforth in this section, all the computations will be made with respect to this system of coordinates defined in a neighborhood of the point $\overline{\xi}$ (which has canonical coordinates $u = 0$).

From now on we will skip the u superscript from $\widetilde{X}_i^u$, recalling that *all our vector fields will be expressed in canonical coordinates* and all functions will be defined in a neighborhood of the origin.

We start with the following:

Lemma 4.10 *In the canonical coordinates we have*

$$\sum_{I\in\mathcal{B}} u_I \partial_{u_I} = \sum_{I\in\mathcal{B}} u_I \widetilde{X}_{[I]}. \tag{4.11}$$

Since the above relation will play an essential role in the following, we present its easy proof.

Proof. We start by noting that, if $Y = \sum_{I\in\mathcal{B}} y_I(u) \frac{\partial}{\partial u_I}$ and $Z = \sum_{I\in\mathcal{B}} z_I(u) \frac{\partial}{\partial u_I}$ are two vector fields such that

$$Z(u_J) = Y(u_J) \text{ for any } J \in \mathcal{B},$$

(that is, the vector fields act at the same way on the functions $u \mapsto u_J$) then $y_I(u) = z_I(u)$ for any $I \in \mathcal{B}$, hence $Y = Z$. Therefore, it will be enough to show that

$$\left(\sum_{I\in\mathcal{B}} u_I \widetilde{X}_{[I]}\right)(u_J) = \left(\sum_{I\in\mathcal{B}} u_I \frac{\partial}{\partial u_I}\right)(u_J)$$

that is

$$\left(\sum_{I\in\mathcal{B}} u_I \widetilde{X}_{[I]}\right)(u_J) = u_J.$$

Now, for any vector field Y,

$$Yf(\xi) = \frac{d}{dt}\left(f\left(\exp(tY)\left(\overline{\xi}\right)\right)\right)_{/t=t_0} \text{ where } \xi = \exp(t_0 Y)\left(\overline{\xi}\right).$$

Hence, if $Y = \sum_{I\in\mathcal{B}} u_I \widetilde{X}_{[I]}$, then

$$\left(\sum_{I\in\mathcal{B}} u_I \widetilde{X}_{[I]}\right)(u_J) = \frac{d}{dt}\left(u_J\left(\exp\left(t\sum_{I\in\mathcal{B}} u_I \widetilde{X}_{[I]}\right)\left(\overline{\xi}\right)\right)\right)_{/t=t_0}$$

just by definition of the coordinates u_I

$$= \frac{d}{dt}(tu_J)_{/t=t_0} = u_J.$$

■

Definition 4.11 (Weights) [2] *We assign the weight $|I|$ to the coordinate u_I and the weight $-|I|$ to ∂_{u_I}. In the following we will say that a C^∞ function f has weight $\geqslant \sigma$ if the Taylor expansion of f at the origin does not include terms of the kind $au_{I_1}u_{I_2}\cdots u_{I_k}$ with $a \neq 0$ and $|I_1| + |I_2| + \ldots + |I_k| < \sigma$. A vector field $Y = \sum_{I\in\mathcal{B}} f_I \partial_{u_I}$ has weight $\geqslant \sigma$ if f_I has weight $\geqslant \sigma + |I|$ for every $I \in \mathcal{B}$.*

Note that the weight of a smooth function is always ≥ 0, while the weight of a vector field is $\geq -s$, since $|I| \leq s$ for every $I \in \mathcal{B}$.

We want to stress that the definition of weight relies on the canonical coordinates, therefore it depends on the choice of a particular basis $\mathcal{B}$ of $\mathbb{R}^N$.

The next theorem contains a fundamental piece of information about the vector fields $\widetilde{X}_{[I]}$ expressed in canonical coordinates, which parallels the properties of left invariant homogeneous vector fields on a homogeneous group:

Theorem 4.12 (Weight of a vector field) *For every multiindex I, the vector field $\widetilde{X}_{[I]}$ has weight $\geqslant -|I|$.*

We can now state the approximation theorem for free weighted vector fields:

Theorem 4.13 (Approximation, pointwise version) *Let $Y_1, \ldots, Y_q$ be another system of vector fields defined in a neighborhood of the origin using the canonical coordinates $\{u_I\}_{I\in\mathcal{B}}$ induced by $\widetilde{X}_{[I]}$ and satisfying*

$$\sum_{I\in\mathcal{B}} u_I \partial_{u_I} = \sum_{I\in\mathcal{B}} u_I Y_{[I]}. \tag{4.12}$$

[2] We alert the reader that the above convention about weights of functions and differential operators is the one made in [34], and is *different* from that made in [27] and [44]: in the last two papers, the authors assign positive weight also to derivatives.

Then

$$\widetilde{X}_{[I]} - Y_{[I]} \text{ has weight } \geq 1 - |I|$$

for any multiindex I *with* $|I| \leq s$. *In particular, for* $I = (i)$ *we have that*

$$\widetilde{X}_i - Y_i \text{ has weight } \geq 0 \text{ for } i = 1, 2, ..., q$$

The above two theorems are the hard core of the approximation result, and their proofs are hard computations.

Note that Theorem 4.13 actually contains result which is more general than the original one by Rothschild-Stein, since it allows to approximate the system of vector fields $\widetilde{X}_i$, in a suitable coordinate system, by *any other system of vector fields* Y_i *satisfying* (4.12).

Nevertheless, what makes this fact really useful is the possibility of choosing as approximating vector fields a family of homogeneous left invariant vector fields on a homogeneous group. This requires an abstract construction.

We have q vector fields $\widetilde{X}_1, ..., \widetilde{X}_q$, free up to step s and satisfying Hörmander's condition at step s in some domain U of $\mathbb{R}^N$ where the integer N only depends on the integers q, s:

$$N = \mathcal{N}(q, s).$$

We have the following result:

Theorem 4.14 *Let* $N = \mathcal{N}(q, s)$. *There exist in* $\mathbb{R}^N$ *a system of smooth vector fields* $Y_1, ..., Y_q$ *and a structure of Carnot group* $\mathbb{G}$ *such that:*

- *(i) the vector fields* $Y_1, ..., Y_q$ *are free up to step* s *in* $\mathbb{R}^N$ *and the vectors* $\left\{ \left(Y_{[I]} \right)_u \right\}_{|I| \leq s}$ *span* $\mathbb{R}^N$ *at any point* u *of the space;*
- *(ii) the* $Y_{[I]}$*'s are left invariant and homogeneous of degree* $|I|$ *with respect to the dilations in* $\mathbb{G}$*;*
- *(iii) if* $I \in \mathcal{B}$ *the vector fields* $Y_{[I]}$ *at* $u = 0$ *coincide with the local basis associated to the coordinates* u_I*, that is,*

$$\left(Y_{[I]} \right)_0 = \frac{\partial}{\partial u_I};$$

- *(iv) the vector field* $Y_{[I]}$ *satisfy (4.12);*
- *(v) in the group* $\mathbb{G}$*, the inverse* u^{-1} *of an element is just its (Euclidean) opposite* $-u$.

Theorem 4.13 can now be applied choosing the left invariant vector fields $Y_{[I]}$ as the approximating system. The map $u = \Theta(\eta, \xi)$ can now be regarded as a diffeomorphism from a neighborhood of η onto a neighborhood of 0 *in the group* $\mathbb{G}$. In other words, $\Theta(\eta,\xi)$ is an element of the group $\mathbb{G}$, and one has:

$$\Theta(\xi,\eta) = -\Theta(\eta,\xi) = \Theta(\eta,\xi)^{-1}.$$

Let us give also an idea of how Theorem 4.14 is proved. As anticipated in §2.3, this requires an abstract construction.

Let $\mathfrak{g}_q$ be the (abstractly defined) free Lie algebra on q generators. For an integer $s > 1$, let $\mathfrak{h}_s$ be the ideal spanned by the commutators of length at least $s+1$. Then we have:

Proposition 4.15 *The quotient* $\mathfrak{g}_{q,s} = \mathfrak{g}_q/\mathfrak{h}_s$ *is a finite dimensional stratified nilpotent Lie algebra of step* s*. More precisely if we let* W_1 *be the span of the* q *generators, and* W_l *be the span of commutators of length* l*, then we have*

$$\mathfrak{g}_{q,s} = \bigoplus_{k=1}^{s} W_k.$$

Moreover if $k+\ell \leqslant s$ *then*

$$[W_k, W_\ell] = W_{k+\ell}$$

while for $k+\ell > s$

$$[W_k, W_\ell] = 0.$$

Definition 4.16 *The Lie algebra*

$$\mathfrak{g}_{q,s} = \mathfrak{g}_q/\mathfrak{h}_s$$

is called the free nilpotent Lie algebra of step s *with* q *generators.*

On this Lie algebra we can natural define a family of dilations, as follows.

Definition 4.17 *Let* $a \in \mathfrak{g}_{q,s}$ *and write*

$$a = \sum_{k=1}^{s} w_k$$

with $w_k \in W_k$*. For every* $\lambda > 0$ *let us define the dilation operators* D_λ

$$D_\lambda : \mathfrak{g}_{q,s} \to \mathfrak{g}_{q,s}$$
$$D_\lambda(a) = \sum_{k=1}^{s} \lambda^k w_k.$$

Proposition 4.18 *The dilations* D_λ *are automorphisms of the Lie algebra* $\mathfrak{g}_{q,s}$*: for every* $a_1, a_2 \in \mathfrak{g}_{q,s}$ *and* $c_1, c_2 \in \mathbb{R}$*, we have*

$$D_\lambda(c_1a_1 + c_2a_2) = c_1D_\lambda(a_1) + c_2D_\lambda(a_2) \tag{4.13}$$

and

$$[D_\lambda(a_1), D_\lambda(a_2)] = D_\lambda([a_1,a_2]). \tag{4.14}$$

In the next step we would like to see $\mathfrak{g}_{q,s}$ as the Lie algebra of a homogeneous Lie group. In order to understand the problem and identify the candidate algebraic structure of the Lie group we are looking for, let us reverse our reasoning for a moment. Assume *we do have* a homogeneous group $\mathbb{G}$ with group operation $\diamond$, let $\mathfrak{g}$ be its Lie algebra and consider the exponential map

$$\begin{aligned} &\mathrm{Exp} : \mathfrak{g} \to \mathbb{G} \\ &\mathrm{Exp} : X \mapsto \mathrm{Exp}\,(X) = \exp\,(X)\,(0) \end{aligned}$$

Let $X, Y \in \mathfrak{g}$. It is easy to prove that

$$\exp\,(Y)\,(x) = x \diamond \mathrm{Exp}\,(Y)\,.$$

Hence for $x = \mathrm{Exp}\,(X) = \exp\,(X)\,(0)$ we have

$$\exp\,(Y) \exp\,(X)\,(0) = \mathrm{Exp}\,(X) \diamond \mathrm{Exp}\,(Y)\,.$$

Under suitable conditions, by the *Baker-Campbell-Hausdorff formula*, there exists one (and only one) vector field $S\,(X, Y)$, computable in some way starting with X, Y, such that

$$\exp\,(Y) \exp\,(X)\,(0) = \exp\,(S\,(X, Y))\,(0) = \mathrm{Exp}\,(S\,(X, Y))\,. \tag{4.15}$$

Actually, BCH theorem states that:

$$\begin{aligned} S\,(X, Y) = &\ X + Y + \frac{1}{2}\,[X, Y] \\ &+ \frac{1}{12}\,\{[X, [X, Y]] - [Y, [X, Y]]\} \\ &- \frac{1}{48}\,\{[Y, [X, [X, Y]]] + [X, [Y, [X, Y]]]\} + ... \\ &\equiv X + Y + \sum_{k=2}^{+\infty} C_k\,(X, Y) \end{aligned}$$

(but the explicit form of this expression is not important).

If we define

$$X \circ Y = S\,(X, Y)\,,$$

then it can be proved that this $\circ$ is a group operation.

Moreover, by (4.15) we have

$$\mathrm{Exp}\,(X \circ Y) = \mathrm{Exp}\,(X) \diamond \mathrm{Exp}\,(Y)\,,$$

so that the mapping

$$\mathrm{Exp} : (\mathfrak{g}, \circ) \to (\mathbb{G}, \diamond)$$

becomes a Lie group isomorphism.

We can now come back to our real situation, where $(\mathbb{G}, \diamond)$ does not exist yet, and the identity (4.15) cannot be used. Nevertheless, the group $(\mathfrak{g}_{q,s}, \circ)$ actually exists and by the previous reasoning should be isomorphic to the desired group $(\mathbb{G}, \diamond)$. Hence, the path is now drawn: starting with the group structure $(\mathfrak{g}_{q,s}, \circ)$ in $\mathbb{R}^M$, we have to check that this is actually a homogeneous group, we have to construct its Lie algebra and to check that it is isomorphic to $\mathfrak{g}_{q,s}$.

So, let us start with the free nilpotent Lie algebra $\mathfrak{g}_{q,s}$, which can be identified with $\mathbb{R}^M$. Let

$$S(u,v) = u + v + \sum_{k=2}^{+\infty} C_k(u,v)$$

be the formal series involved in BCH formula. Recall that C_k is a homogeneous Lie polynomial of degree k. In our nilpotent Lie algebra, when $k > s$ and $u, v \in \mathbb{R}^M$ we have $C_k(u,v) = 0$ so that we can define

$$u \circ v = S(u,v) = u + v + \sum_{k=2}^{s} C_k(u,v).$$

Remark 4.19 *Let $S(u,v) = (S_1(u,v), S_2(u,v), \ldots, S_M(u,v))$. It is a simple computation to check that every $S_i(u,v)$ is actually a polynomial of degree at most s in the variables u_i, v_i.*

Proposition 4.20 *The binary operation*

$$u \circ v = S(u,v) = u + v + \sum_{k=2}^{s} C_k(u,v)$$

and the dilations D_λ induce in $\mathbb{R}^M$ a structure of homogeneous group $\mathbb{G}$. The inverse of a element $u \in \mathbb{G}$ is the opposite $-u$.

References

[1] S. Biagi, A. Bonfiglioli: The existence of a global fundamental solution for homogeneous Hörmander operators via a global lifting method. Proc. Lond. Math. Soc. (3) 114 (2017), no. 5, 855–889.

[2] S. Biagi, A. Bonfiglioli: Global Heat Kernels for Parabolic Homogeneous Hörmander Operators. To appear on Israel J. of Math. arxiv.org/abs/1910.09907

[3] S. Biagi, A. Bonfiglioli, M. Bramanti: Global estimates for the fundamental solution of homogeneous Hörmander sums of squares. Journal of Math. Anal. and Appl. Volume 498, Issue 1, 1 June 2021, 124935.

[4] S. Biagi, M. Bramanti: Global estimates for the heat kernels of homogeneous Hörmander sums of squares. (2020). Potential Analysis. Published online 8 november 2021.

[5] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: Stratified Lie groups and potential theory for their sub-Laplacians. Springer Monographs in Mathematics. Springer, Berlin, 2007.

[6] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: Uniform Gaussian estimates of the fundamental solutions for heat operators on Carnot groups, Adv. Differential Equations, 7 (2002), 1153-1192.

[7] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: Fundamental solutions for non-divergence form operators on stratified groups. Trans. Amer. Math. Soc. 356 (2004), no. 7, 2709-2737

[8] A. Bonfiglioli, F. Uguzzoni: Families of diffeomorphic sub-Laplacians and free Carnot groups. Forum Math. 16 (2004), no. 3, 403-415.

[9] A. Bonfiglioli, F. Uguzzoni: A note on lifting of Carnot groups. Rev. Mat. Iberoamericana 21 (2005), no. 3, 1013–1035.

[10] A. Bonfiglioli, F. Uguzzoni: Harnack inequality for non-divergence form operators on stratified groups, Trans. Amer. Math. Soc. 359 (2007), 2463-2481.

[11] M. Bramanti: An invitation to hypoelliptic operators and Hörmander's vector fields. SpringerBriefs in Mathematics. Springer, Cham, 2014. xii+150 pp.

[12] M. Bramanti. On the proof of Hörmander's hypoellipticity theorem. Le Matematiche, special issue "New trends in PDEs" (Catania, May 29-30th 2018). Vol 75 No 1 (2020), 3-26.

[13] M. Bramanti, L. Brandolini: Hörmander operators. World Scientific (2023), pp.XXVIII+693.

[14] M. Bramanti-L. Brandolini: L^p-estimates for uniformly hypoelliptic operators with discontinuous coefficients on homogeneous groups. Rend. Sem. Mat. dell'Univ. e del Politec. di Torino. Vol. 58, 4 (2000), 389-433.

[15] M. Bramanti, L. Brandolini: L^p-estimates for nonvariational hypoelliptic operators with VMO coefficients. Trans. Amer. Math. Soc. 352 (2000), no. 2, 781-822.

[16] M. Bramanti, L. Brandolini, E. Lanconelli, F. Uguzzoni: Non-divergence equations structured on Hörmander vector fields: heat kernels and Harnack inequalities. Memoirs of the AMS 204 (2010), no. 961, 1-136.

[17] M. Bramanti, L. Brandolini, M. Manfredini, M. Pedroni: Fundamental solutions and local solvability for nonsmooth Hörmander operators. Memoirs of the AMS, vol. 249, no. 1182. (2017).

[18] M. Bramanti, L. Brandolini, M. Pedroni: On the lifting and approximation theorem for nonsmooth vector fields. Indiana University Mathematics Journal. Issue 6 Volume 59 (2010), 1889–1934.

[19] M. Bramanti, L. Brandolini, M. Pedroni: Basic properties of nonsmooth Hörmander's vector fields and Poincaré's inequality. Forum Mathematicum. Volume 25, Issue 4, Pages 703–769 (2013).

[20] M. Bramanti, M. Zhu: $L^p$and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift. Analysis & PDE 6-8 (2013), 1793-1855.

[21] A. P. Calderón, A. Zygmund: On the existence of certain singular integrals. Acta Math. 88, (1952). 85–139.

[22] A. P. Calderón, A. Zygmund. On singular integrals. Amer. J. of Math. 78 (1956), 249-271.

[23] A. P. Calderón, A. Zygmund: Singular integral operators and differential equations. Amer. J. Math. 79 1957 901–921.

[24] M. Christ, A. Nagel, E. M. Stein, S. Wainger: Singular and maximal Radon transforms: analysis and geometry. Ann. of Math. (2) 150 (1999), no. 2, 489–577.

[25] R. R. Coifman, G. Weiss: Analyse harmonique non-commutative sur certains espaces homogènes. Étude de certaines intégrales singulières. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York, 1971.

[26] C. Fefferman, A. Sánchez-Calle: Fundamental solutions for second order subelliptic operators. Ann. of Math. (2) 124 (1986), no. 2, 247–272.

[27] G. B. Folland: Subelliptic estimates and function spaces on nilpotent Lie groups. Ark. Mat. 13 (1975), no. 2, 161–207.

[28] G. B. Folland: On the Rothschild-Stein lifting theorem. Comm. Partial Differential Equations 2 (1977), no. 2, 165–191.

[29] G. B. Folland: Applications of analysis on nilpotent groups to partial differential equations. Bull. Amer. Math. Soc. 83 (1977), no. 5, 912–930.

[30] G. B. Folland, E. M. Stein: Parametrices and estimates for the $\overline{\partial}_b$ complex on strongly pseudoconvex boundaries. Bull. Amer. Math. Soc. 80 (1974), 253–258.

[31] G. B. Folland, E. M. Stein: Estimates for the $\overline{\partial}_b$ complex and analysis on the Heisenberg group. Comm. Pure Appl. Math. 27 (1974), 429–522.

[32] R. W. Goodman: Lifting vector fields to nilpotent Lie groups. J. Math. Pures Appl. (9) 57 (1978), no. 1, 77–85.

[33] L. Hörmander: Hypoelliptic second order differential equations. Acta Math. 119 (1967), 147-171.

[34] L. Hörmander, A. Melin: Free systems of vector fields. Ark. Mat. 16 (1978), no. 1, 83–88.

[35] D. Jerison: The Poincaré inequality for vector fields satisfying Hörmander's condition. Duke Math. J. 53 (1986), no. 2, 503–523.

[36] A. W. Knapp, E. M. Stein: Intertwining operators for semisimple groups. Ann. of Math. (2) 93 (1971), 489–578.

[37] J. J. Kohn: Pseudo-differential operators and hypoellipticity. Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971), pp. 61–69. Amer. Math. Soc., Providence, R.I., 1973.

[38] A. N. Kolmogorov: Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). Ann. of Math. (2) 35 (1934), no. 1, 116–117.

[39] E. Lanconelli, D. Morbidelli: On the Poincaré inequality for vector fields. Ark. Mat. 38 (2000), no. 2, 327–342.

[40] A. Montanari, D. Morbidelli: Nonsmooth Hörmander vector fields and their control balls. Trans. Amer. Math. Soc 364, (2012), 2339–2375.

[41] D. Mumford: Elastica and computer vision. Algebraic geometry and its applications (West Lafayette, IN, 1990), 491–506, Springer, New York, 1994.

[42] A. Nagel, E. M. Stein, S. Wainger: Balls and metrics defined by vector fields. I. Basic properties. Acta Math. 155 (1985), no. 1-2, 103–147.

[43] O. A. Oleĭnik, E. V. Radkevič: Second order equations with nonnegative characteristic form. Plenum Press, New York-London, 1973. (Translated from the original version in Russian, 1969).

[44] L. P. Rothschild, E. M. Stein: Hypoelliptic differential operators and nilpotent groups. Acta Math. 137 (1976), no. 3-4, 247–320.

[45] A. Sánchez-Calle: Fundamental solutions and geometry of the sum of squares of vector fields. Invent. Math. 78 (1984), no. 1, 143–160.

[46] E. M. Stein: Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

Marco Bramanti
Dipartimento di Matematica
Politecnico di Milano
Via Bonardi 9. 20133 Milano (Italy)

https://bramanti.faculty.polimi.it
marco.bramanti@polimi.it